**FINAL REPORT**

**TEAM NAME**: SW\_2016

**TEAM MEMBERS**: Indhumadhi Suryanarayanan

Sindhujaa Ranganathan Chandra Babu

**PROJECT TITLE**: Bank Activities and Fails.

**PROJECT TYPE:** LOGD

**INTRODUCTION:**

This is a LOGD Project demonstrating integration involving semantic web technologies. We use the framework and data provided at the Linking Open Government Data web site as the foundation for our project.

The project involves understanding the semantic representation of data sources and use f triples from two different data sets to represent knowledge.

Taking the above problem into consideration, we have decided to analyze the scenario whether the assets allocated to a particular bank plays any factor in deciding the failure of the bank. If the results are positive, if we can figure a pattern between the assets belonging to a particular state in a particular range of time and number of closed banks in a state for the same period of time, we can monitor the assets assignment and accordingly assign the right amount for the future.

**TARGET AUDIENCE:**

The **depositors, creditors, customers and employees** are all direct victims of

Bank Failure. Basically anyone who is part of any kind of interaction involving a Bank are strongly influenced by the state of the bank. Any kind of **investors** would also be interested in the activities of the bank.

The focus of this project would be more relatable for anyone from **the government body** or any other **public offices** concerned with **finances** of the state and country to keep track of banking relate activities.

**DESCRIPTION OF DATA SOURCES:**

**DATASETS USED:**

**ID: 1193**

**TRIPLES COUNT: 331**

**PUBLISHER:** Department of the Treasury

**DETAILS:** The OCC's quarterly report on bank derivatives activities and trading revenues is based on Call Report information provided by all insured U.S. commercial banks and trust companies, as well as on other published financial data.

**ID: 1374**

**TRIPLES COUNT: 1,826**

**PUBLISHER:** Federal Deposit Insurance Corporation

**DETAILS:** The FDIC is often appointed as receiver for failed banks. This list includes banks which have failed since October 1, 2000.

**SAMPLE TRIPLES FOR DATASET 1374**

<rdf:Description rdf:about="#entry202">

<updated\_date>23-Mar-10</updated\_date>

<closing\_date>6-Feb-09</closing\_date>

<cert>57017</cert>

<state>GA</state>

<city>McDonough</city>

<bank\_name>FirstBank Financial Services</bank\_name>

<rdf:type

rdf:resource="http://data-gov.tw.rpi.edu/2009/data-gov-twc.rdf#Da

taEntry"/>

</rdf:Description>

**SAMPLE TRIPLES FOR DATASET 1193**

<rdf:Description rdf:about="#entry23">

<total\_futures\_exch\_tr>0</total\_futures\_exch\_tr>

<rank>23</rank>

<total\_forwards\_otc>352</total\_forwards\_otc>

<total\_credit\_derivatives\_otc>251</total\_credit\_derivatives\_otc>

<bank\_name>TD BANK NATIONAL ASSN</bank\_name>

<rdf:type

rdf:resource="http://data-gov.tw.rpi.edu/2009/data-gov-twc.rdf#DataEntry"/>

<total\_swaps\_otc>18367</total\_swaps\_otc>

<total\_options\_otc>9717</total\_options\_otc>

<spot\_fx>21</spot\_fx>

<total\_derivatives>28687</total\_derivatives>

<total\_options\_exch\_tr>0</total\_options\_exch\_tr>

<state>DE</state>

<total\_assets>104413</total\_assets>

</rdf:Description>

**DATA INTEGRATION:**

The two data sets could be integrated through the common field **“state”.** Since data set 1193 is available for only the period of APRIL 2009 – June 2009, the integration is done only on the filtered results of Dataset 1374 during the same time period.

We have used various aggregate functions to compute total count. The highest of the **total\_assets** and failed bank count is the major area of concentration. We have also shown the trend of bank failures over the years for selected cities.

DATA PRODUCT RESULTS:

**SPARQL QUERY FOR THE FAILED BANK TREND:**

PREFIX d1374: <http://data-gov.tw.rpi.edu/vocab/p/1374/>

SELECT ?year ?state (count(?bank) as ?count) WHERE {

?subject d1374:state ?state.

?subject d1374:bank\_name ?bank.

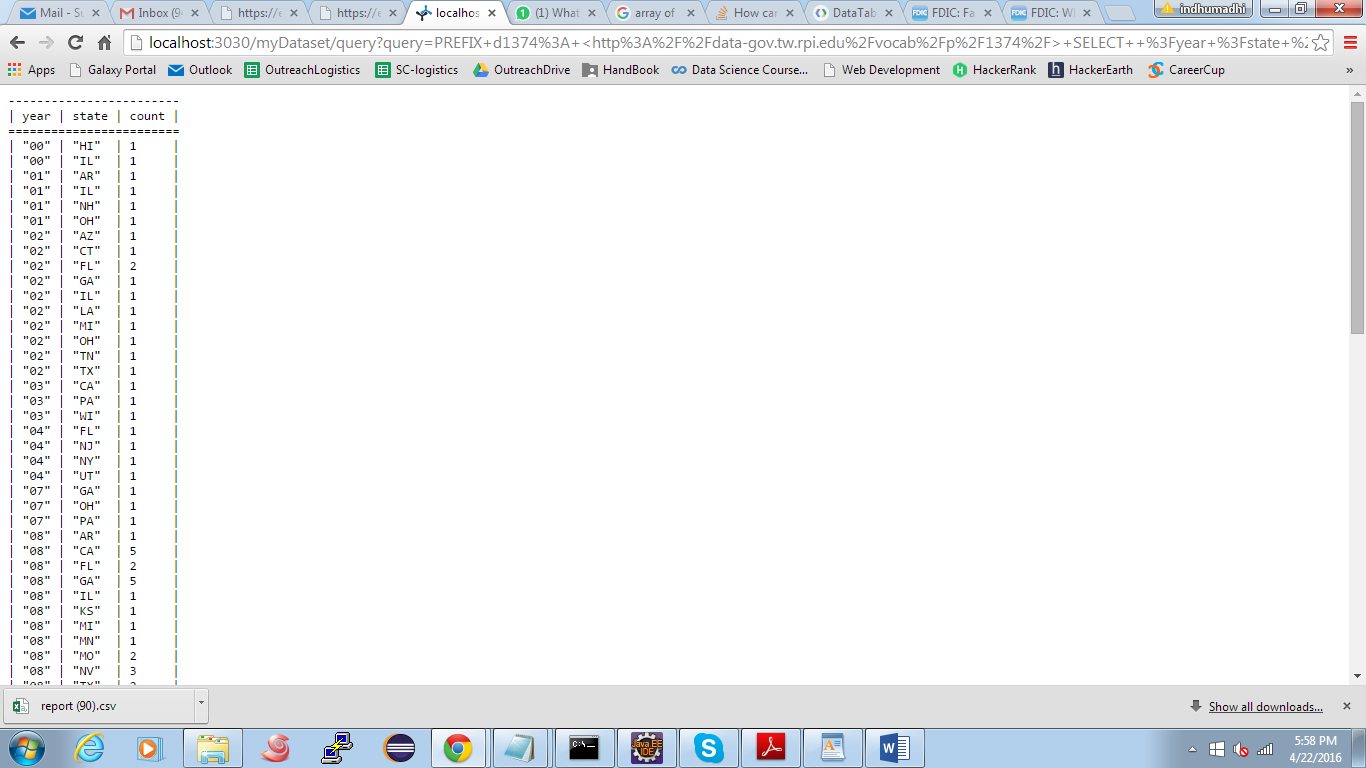
?subject d1374:closing\_date ?closingDate.

bind(if (STRLEN( STR(?closingDate)) =9, SUBSTR(STR(?closingDate), 8, 2),

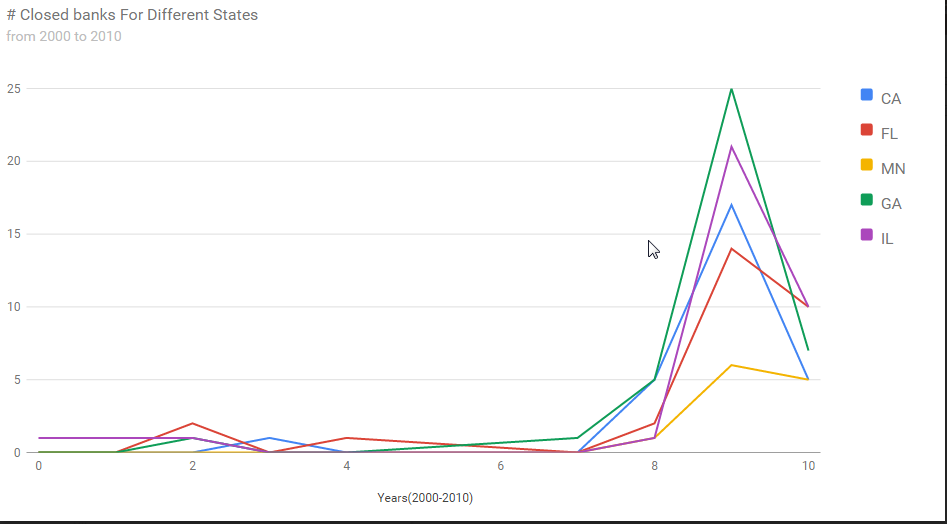
SUBSTR(STR(?closingDate), 7, 2)) as ?year )

} GROUP BY ?year ?state ORDER BY ?year ?state

**SPARQL OUTPUT THROUGH FUSEKI:**



**GRAPH PLOTTED BASED ON THE ABOVE QUERY RESULTS:**



**SPARQL QUERY FOR THE TOTAL\_ASSETS AND FAILED BANK COUNT COMPARISON:**

PREFIX xsd: [http://www.w3.org/2001/XMLSchema#](http://www.w3.org/2001/XMLSchema)

select ?o (COUNT(?o) AS ?closed\_count) (max(xsd:integer(?a)) as ?asset) where {

?s <http://data-gov.tw.rpi.edu/vocab/p/1374/state> ?o .

?s <http://data-gov.tw.rpi.edu/vocab/p/1374/closing\_date> ?d.

?t <http://data-gov.tw.rpi.edu/vocab/p/1193/state> ?o .

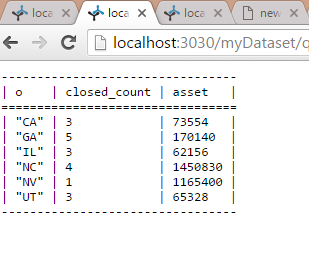
?t <http://data-gov.tw.rpi.edu/vocab/p/1193/total\_assets> ?a.

FILTER (regex(str(?d), "Apr") || regex(str(?d), "May")||regex(str(?d), "Jun")).

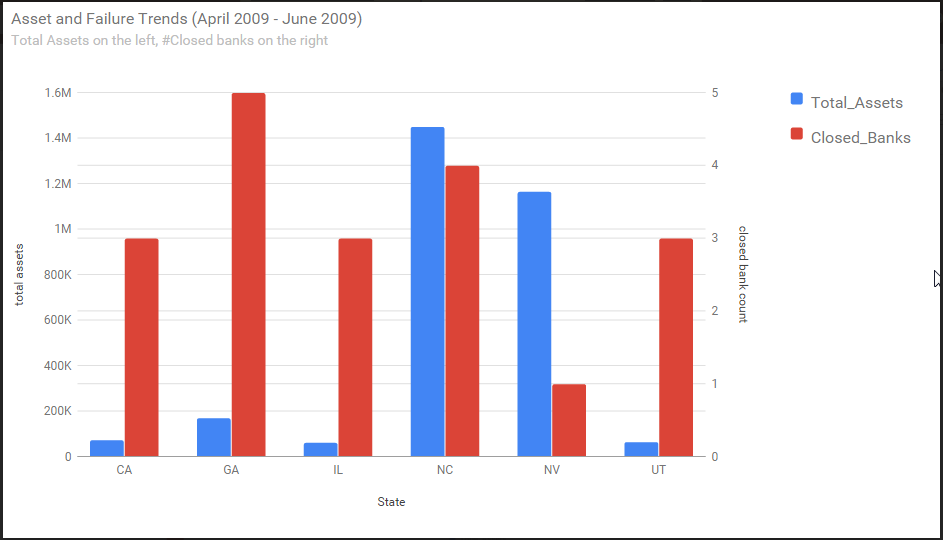
FILTER (regex(substr(?d, strlen(?d)-1, strlen(?d)), "09")).

} group by ?o order by ?o

**SPARQL OUTPUT THROUGH FUSEKI:**



**GRAPH PLOTTED BASED ON THE ABOVE QUERY RESULTS:**



**SUMMARY:**

**CRITICAL INFORMATION:**

The following information can be derived from the graphs:

**\***The Bank failure count saw a huge rise in the year 2009 across few major states in the country which could have been the result of some financial instability.

**\***Through which it can be concluded that in 2009 the bank failures during that period were not individual failures but a result of a national economy change

**\***We aimed at finding out a pattern between the highest asset values of banks in a state and the failure rate at that state. Looks like the financial status of the state does not reflect the failure of the banks.

**OUTSTANDING ISSUES:** There are no outstanding issues. We have produced the results that were promised.

**THINGS THAT COULD HAVE BEEN DONE:**

**\***If data was present over a longer time period and for comprehensive list of banks we could have derived lot more details from the graphs.

**\***We could not integrate data based on the “bank\_name” since the list of banks in both the data sets were mutually exclusive.

**\***While one of the dataset contained data for the years 2000-2010, the other had data only for the second quarter of 2009. A wider data could have resulted in exhaustive inferences about the banking scenario from the graphs.

**\***In future we could analyze the attributes like the “total\_derivatives” to see if they influence the banks standing on the financial front.