

Final Report

Project Title: Crime Analysis of Baltimore City

Team Name: Web Crawlers

Team Members:

Jithendhiralal Gujuluva Ramlal – jgr150030@utdallas.edu

Ramakrishnan Sathyavageeswaran – rxs142530@utdallas.edu

Type of Project: Custom Project

1. Introduction:

- A detailed crime analysis of Baltimore city can occur at various levels, including tactical, operational, and strategic. We are studying crime reports, arrests reports, and police calls for service to identify emerging patterns, series, and trends as quickly as possible. We analyze these phenomena for all relevant factors, predict future occurrences, and issue reports, and alerts to their agencies.
- Rdf123 is used to convert the csv data to RDF raw data.
- Apache-Jena-Fuseki-2.3.0 is used to host the Datasets.
- JavaScript, AJAX, HTML, CSS, jQuery, Google Visualization API and Bootstrap are used to provide the end results to our webpage.
- Google maps, Bar chart, Column chart are the Google Visualization API's used to display the results.
- Google maps depict the entire Baltimore city with the red markers displaying what kind of crime it is and blue markers depicting the CCTV location with their name.
- Clicking at red markers in the google maps displays the kind of crime and for blue markers it displays the CCTV name.
- Drop down list is used to display crime based on the neighborhood in Baltimore.
- Column chart is used to displays the neighborhood vs crime count.
- Bar chart is used to displays the unemployment rate vs neighborhood.
- All these information's are displayed in a webpage.

2. Target Audience:

- Our project combines the crime and its related datasets and highlight the Baltimore's crime rate which helps us in identify emerging patterns, series, and trends as quickly as possible.
- Our target audience are Baltimore police department and City residents.
- BPD will find the results helpful in predicting future occurrences, and issuing reports in crime.
- With the help of these analysis, we can even predict which of these neighborhood are safe to be a residential area.

3. Description of Data Sources:

All below datasets are taken from: <https://data.baltimorecity.gov/>

- **Recorded Crimes by Region FY 2014:**

- This dataset provides the crime code, description, crime date, geo location (with latitude and longitude) and neighborhood.
- Kinds of recorded crimes are Robbery, Rape, Prostitution, Carjacking, aggravated assault.
- **Unemployment rate by Region:**
 - This raw RDF had various information about every neighborhood in the city. From all the information, we need only unemployment rate.
 - After data formatting and data cleaning done using **Google refine**, we represent only the neighborhood and their unemployment rate from this dataset.
- **CCTV location data:**
 - Provides information about the CCTV location with latitude and longitude, their names and their district.

4. Data Integration:

- The three Datasets are integrated using common predicate named 'neighborhood'.
- CCTV location data is mainly used for showcasing it in google maps. And we derived a relationship between crime and unemployment of that area.
- We hosted SPARQL endpoints for all three Datasets using Jena-Fuseki server.
- SPARQL query used

Query for unemployment rate :

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ds: <http://data.baltimorecity.gov/resource/_8hgm-7t56/>
PREFIX ts: <http://data.baltimorecity.gov/resource/_59fg-ary5/>
PREFIX dsbase: <http://data.baltimorecity.gov/resource/>
PREFIX geo: <http://www.w3.org/2003/01/geo/wgs84_pos#>
SELECT DISTINCT ?neighborhood1 ?unemp_rate
WHERE {
  ?element rdf:type dsbase:_8hgm-7t56;
  ds:neighborhood ?neighborhood1;
  ds:unempr13 ?unemp_rate.
  ?element1 rdf:type dsbase:_59fg-ary5;
  ts:neighborhood ?neighborhood;
  filter contains(?neighborhood,?neighborhood1)
} ORDER BY ASC(?neighborhood1)
```

Query for Crime count :

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ds: <http://data.baltimorecity.gov/resource/_8hgm-7t56/>
PREFIX ts: <http://data.baltimorecity.gov/resource/_59fg-ary5/>
PREFIX dsbase: <http://data.baltimorecity.gov/resource/>
PREFIX geo: <http://www.w3.org/2003/01/geo/wgs84_pos#>
```

```

SELECT ?neighborhood1 (count(?neighborhood) AS ?count)
WHERE {
?element rdf:type dsbase:_8hgm-7t56;
ds:neighborhood ?neighborhood1;
ds:unempr13 ?unemp_rate.
?element1 rdf:type dsbase:_59fg-ary5;
ts:neighborhood ?neighborhood;
filter contains(?neighborhood,?neighborhood1)
} GROUP BY ?neighborhood1 ORDER BY ASC(?neighborhood1)

```

Query for Unemployment rate neighborhood wise :

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ds: <http://data.baltimorecity.gov/resource/_8hgm-7t56/>
PREFIX ts: <http://data.baltimorecity.gov/resource/_59fg-ary5/>
PREFIX dsbase: <http://data.baltimorecity.gov/resource/>
PREFIX geo: <http://www.w3.org/2003/01/geo/wgs84_pos#>
SELECT ?neighborhood ?unemp_rate ?code ?desc ?lat ?lang
WHERE {
?element rdf:type dsbase:_8hgm-7t56;
ds:neighborhood "Allendale"; -> give the neighborhood here, for example
"Allendale"
ds:unempr13 ?unemp_rate.
?element1 rdf:type dsbase:_59fg-ary5;
ts:neighborhood ?neighborhood;
filter contains(?neighborhood,"Allendale") -> give the neighborhood here
?element1 ts:crimecode ?code;
ts:description ?desc;
geo:lat ?lat;
geo:long ?lang.
}

```

Query for CCTV location :

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ds: <http://data.baltimorecity.gov/resource/_8hgm-7t56/>
PREFIX ts: <http://data.baltimorecity.gov/resource/_59fg-ary5/>
PREFIX cs: <http://data.baltimorecity.gov/resource/hdyb-27ak/>
PREFIX dsbase: <http://data.baltimorecity.gov/resource/>
PREFIX geo: <http://www.w3.org/2003/01/geo/wgs84_pos#>
SELECT ?name ?lat ?long
WHERE {
?element rdf:type dsbase:hdyb-27ak;
cs:name ?name;
cs:location_1 ?blank.
?blank geo:lat ?lat;

```

```
    geo:long ?long.  
}
```

Query for Unemployment rate for all neighborhood :

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>  
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>  
PREFIX ds: <http://data.baltimorecity.gov/resource/_8hgm-7t56/>  
PREFIX ts: <http://data.baltimorecity.gov/resource/_59fg-ary5/>  
PREFIX dsbase: <http://data.baltimorecity.gov/resource/>  
PREFIX geo: <http://www.w3.org/2003/01/geo/wgs84_pos#>  
SELECT ?neighborhood ?unemp_rate ?code ?desc ?lat ?long  
WHERE {  
    ?element rdf:type dsbase:_8hgm-7t56;  
    ds:neighborhood ?neighborhood1;  
    ds:unempr13 ?unemp_rate.  
    ?element1 rdf:type dsbase:_59fg-ary5;  
    ts:neighborhood ?neighborhood;  
    filter contains(?neighborhood,?neighborhood1)  
    ?element1 ts:crimecode ?code;  
    ts:description ?desc;  
    ts:location_1 ?blank.  
    ?blank geo:lat ?lat;  
    geo:long ?long.  
}
```

5. Data Product Results:

The following are the query response screenshots from fuseki which we are extracting the required data to be displayed in the web page. The Output will be in JSON response.

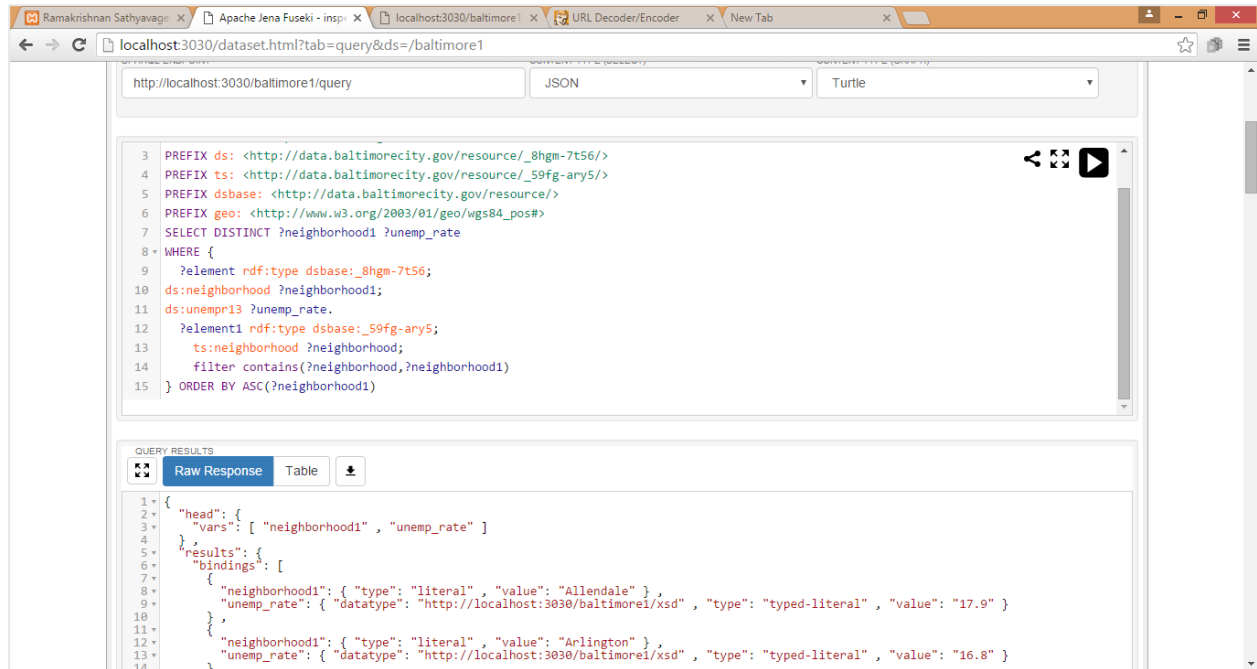


Figure 5.1 Unemployment rate Query

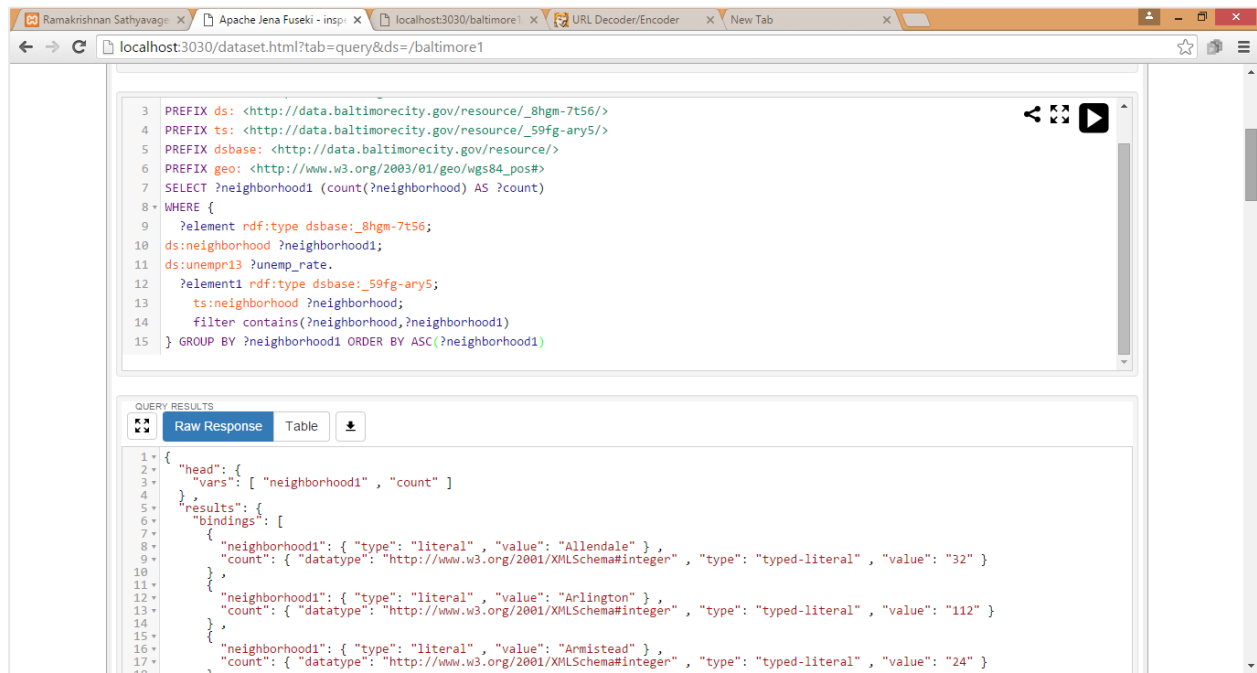


Figure 5.2 Crime Count Query

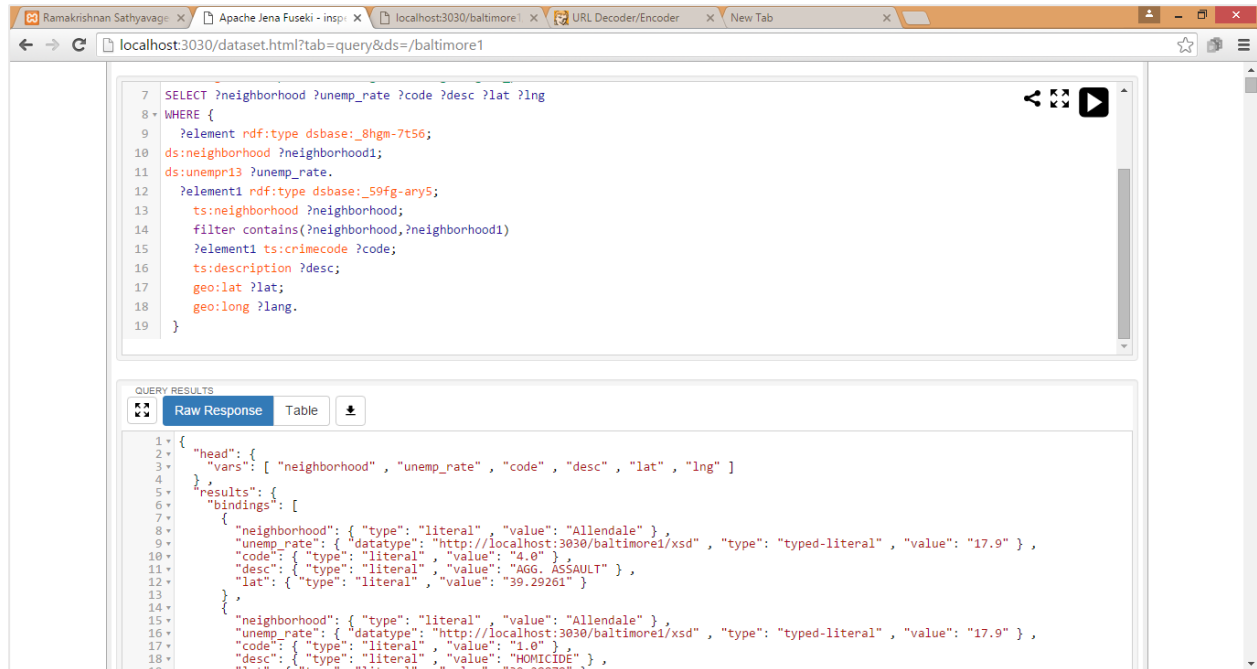


Figure 5.3 All Crime Baltimore crime for Map

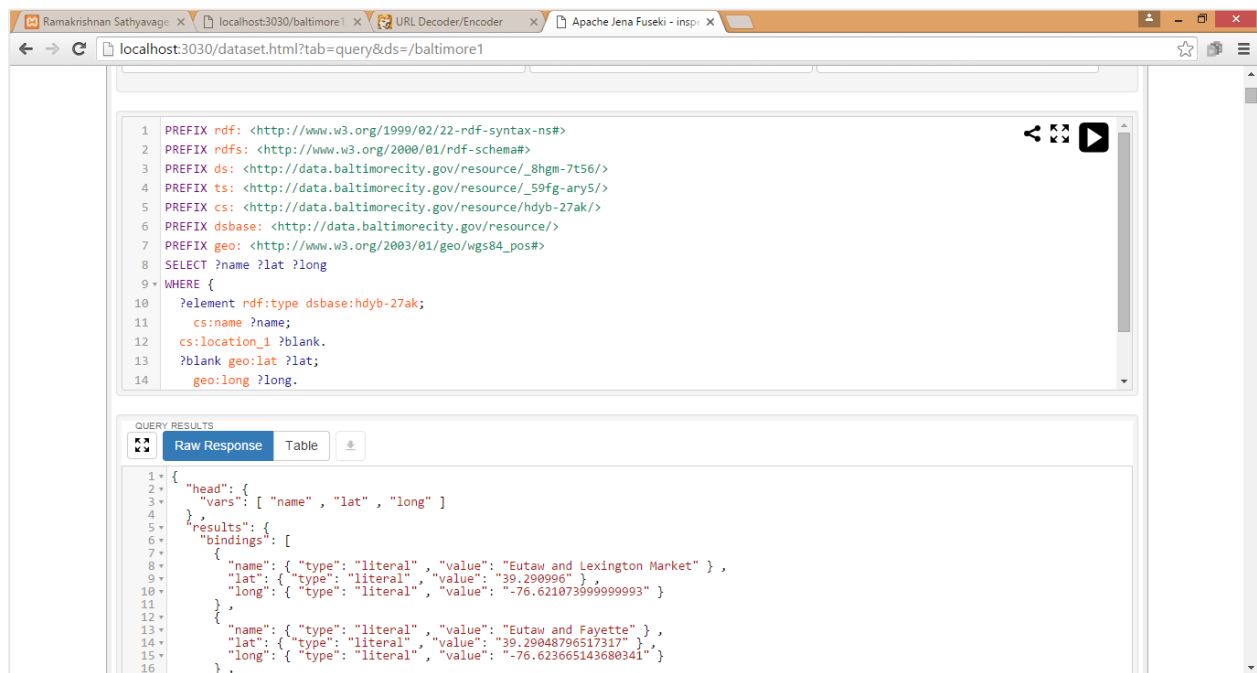


Figure 5.4 CCTV Query response

The following are the screenshots from the web site.

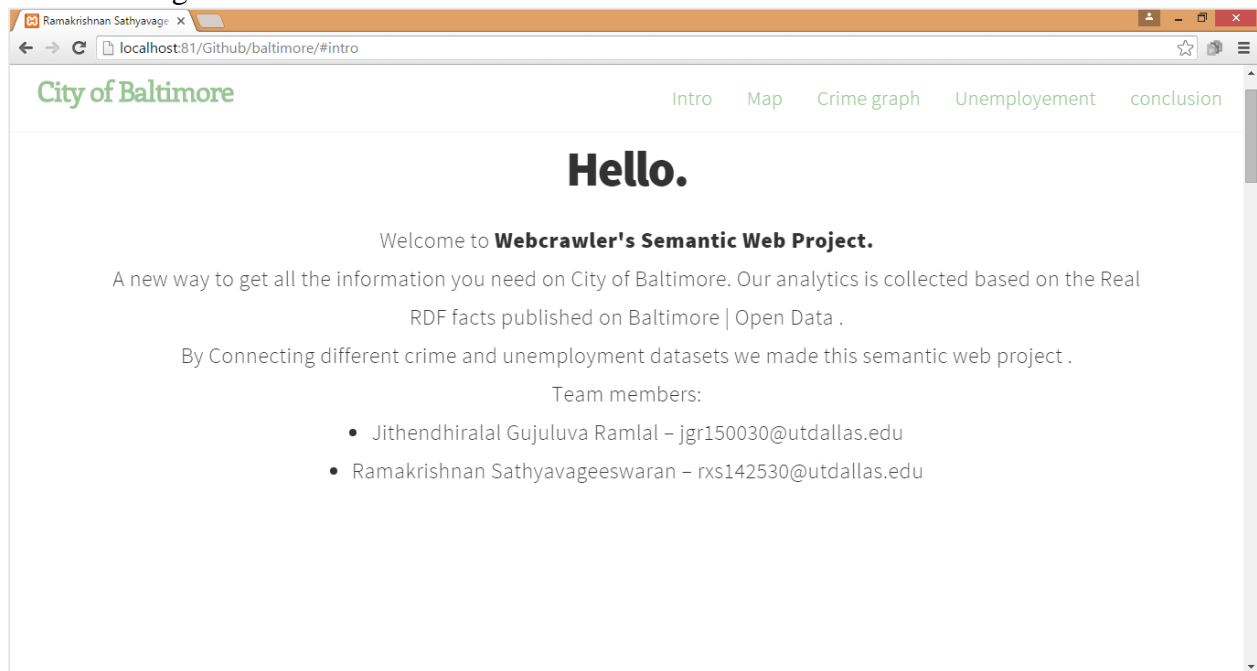


Figure 5.5 Web Page Home screen

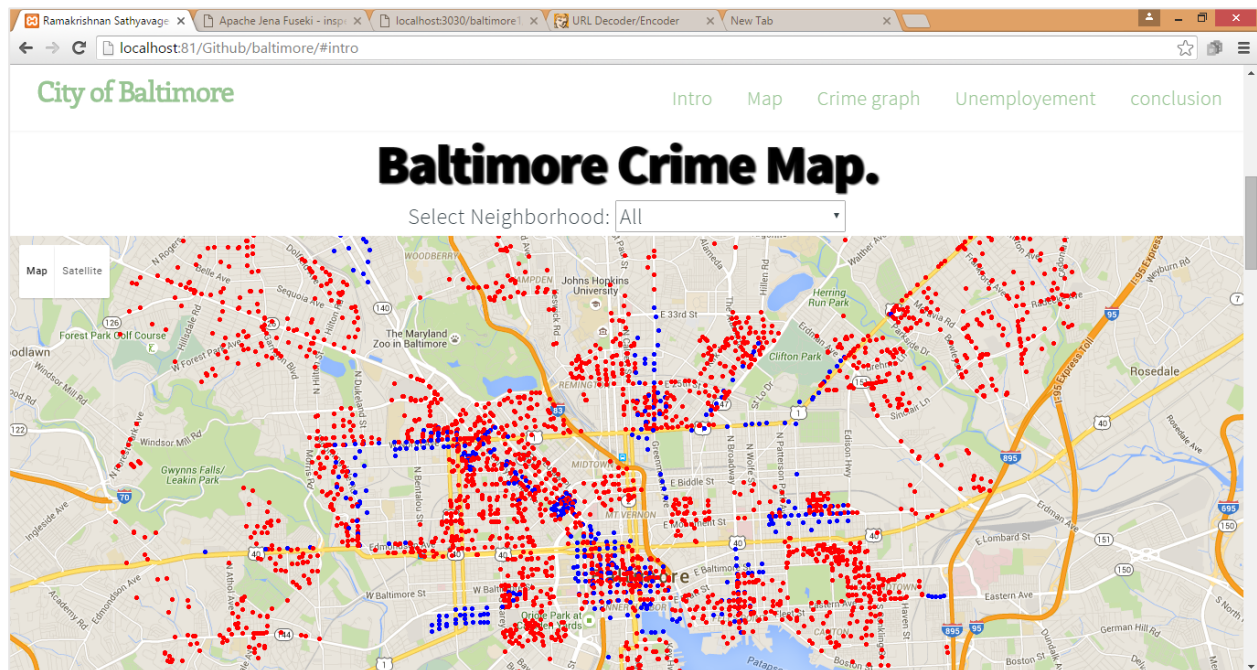


Figure 5.6 Baltimore Crime Map with CCTV location

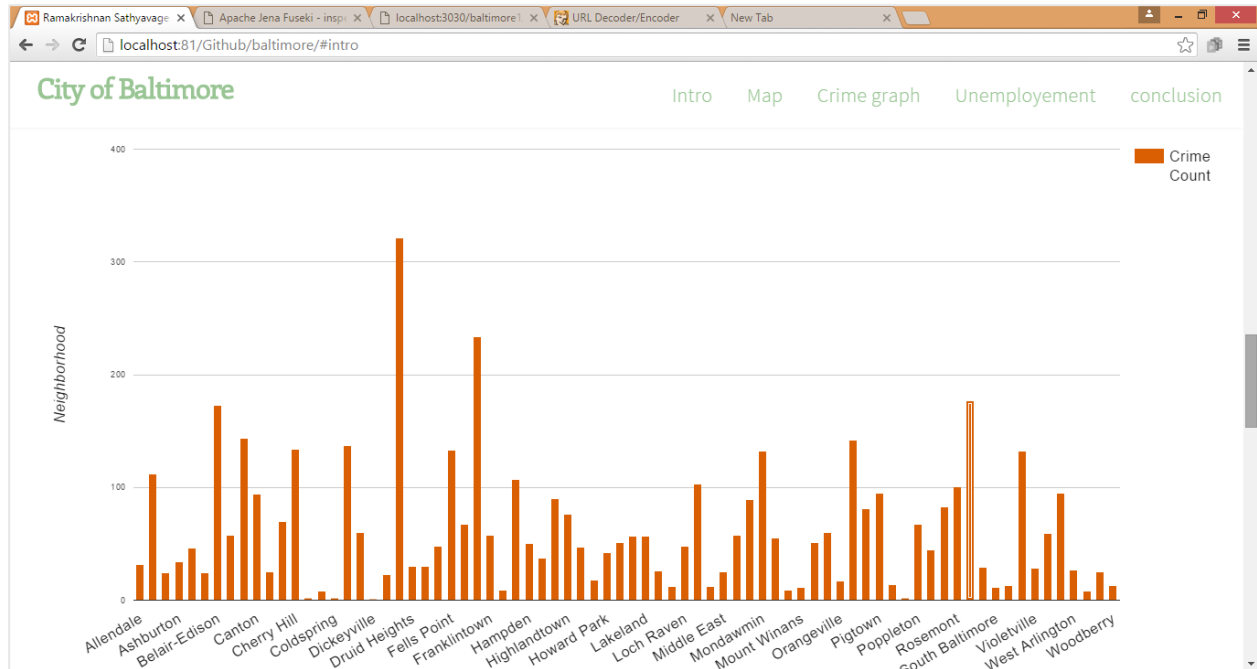


Figure 5.7 Baltimore Crime Count in FY 2014

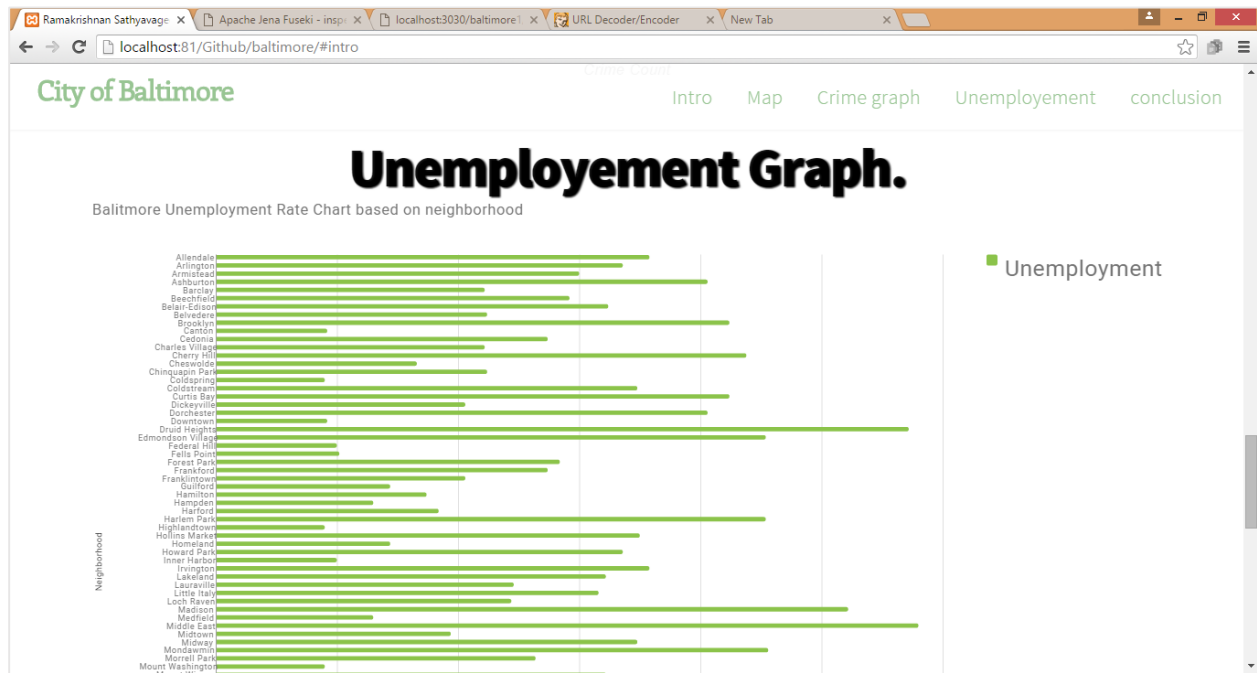


Figure 5.8 Unemployment Graph in FY 2014

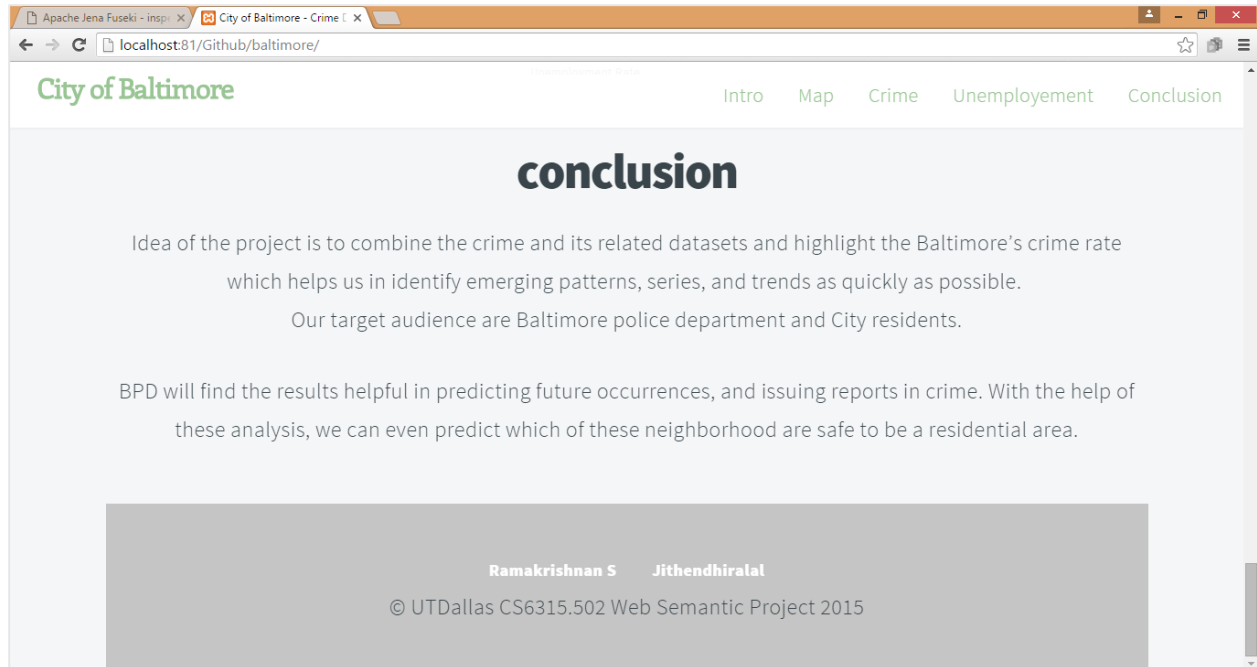


Figure 5.9 Conclusion Page

6. Custom Project Justification:

We started with an idea of crime analysis and then acted upon searching the data. After some ground work, we came to know that Baltimore is one of the city where less data analytics are made on their crime data. We preferred crime analysis on Baltimore city because it's one of the most dangerous cities in USA in terms of crimes.

There are no raw RDF data available in LOGD site for Baltimore city. We constructed RDF file from scratch using the CSV files which we collected from <https://data.baltimorecity.gov/>. Unlike other projects in LOGD, we used maps and charts to provide detailed statistical analysis.

With the gathered dataset, we created a SPARQL endpoint locally using fuseki server.

7. Summary

- Our project combined the crime and its related datasets and highlight the Baltimore's crime rate which helps us in identify emerging patterns, series, and trends as quickly as possible.
- These representations will help the Baltimore police department to control the crime rate and help City residents to choose a safe residential area.