A CUSTOM PROJECT ON

CRIME RATE ANALYSIS (2014&2016)

Project by:

Arun Kumar Kanmani(axk166630)

Nitesh Srivatsav(nxb162630)

Sreenivas Venkitachalam(sxv163530)

Vignesh Vijaykumar(vxv160730)

Table of Contents

1.Introduction

2.Target Audience

3.Description of Data Sources

4.Data Integration

5.Data Product Results

6. Justification

7. Summary

1.Introduction

In a city like Los Angeles which is the second most populous city in the United States crimes are expected to be recurrent. In fact, there are more than 500 crimes reported every day to the Los Angeles Police Department. Our project aims to study the crime rate over the years in the entire city of L.A.

2.Target Audience

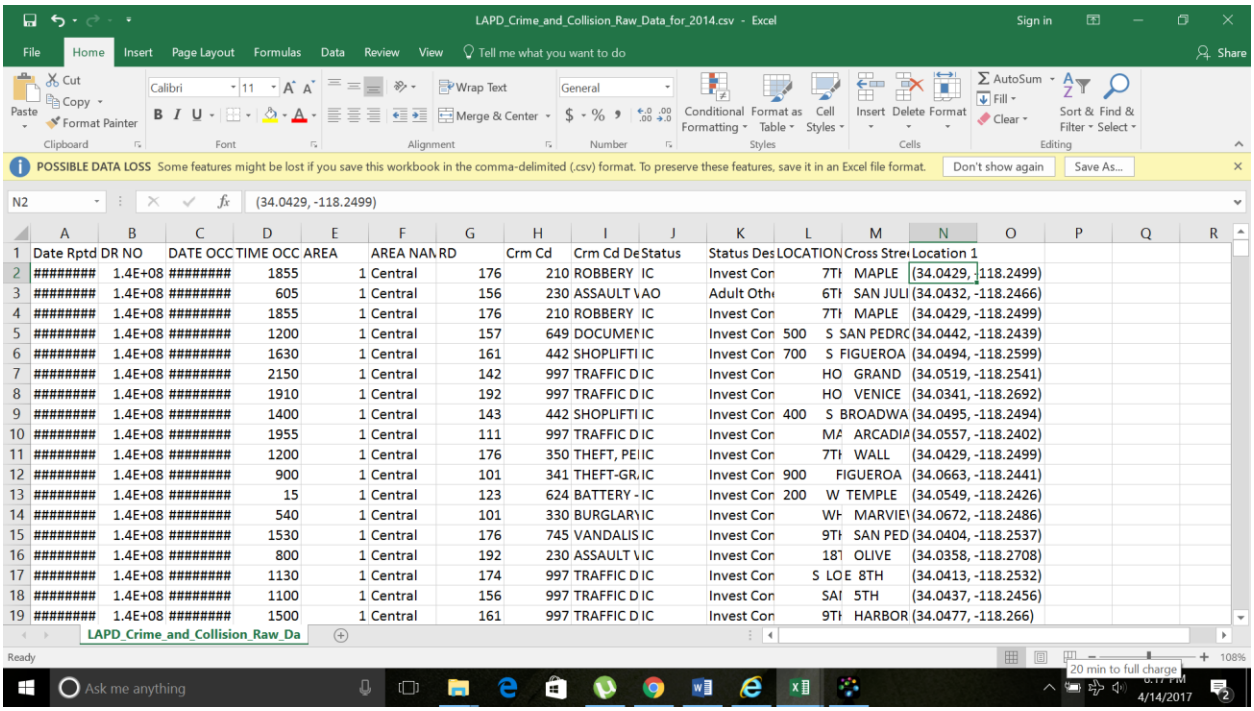
Our project extracts information from two datasets and displays lucent information on the different types of crimes occurring in the city, the crime rates in different areas of the city and the overall efficiency of the Los Angeles Police Department (L.A.P.D).

This project aims to help the L.A.P.D understand the crime occurrences in their city in a better and easier manner using semantic web technologies and google visualization. If the police know areas and types of crimes occurring in them, they would surely be better off.

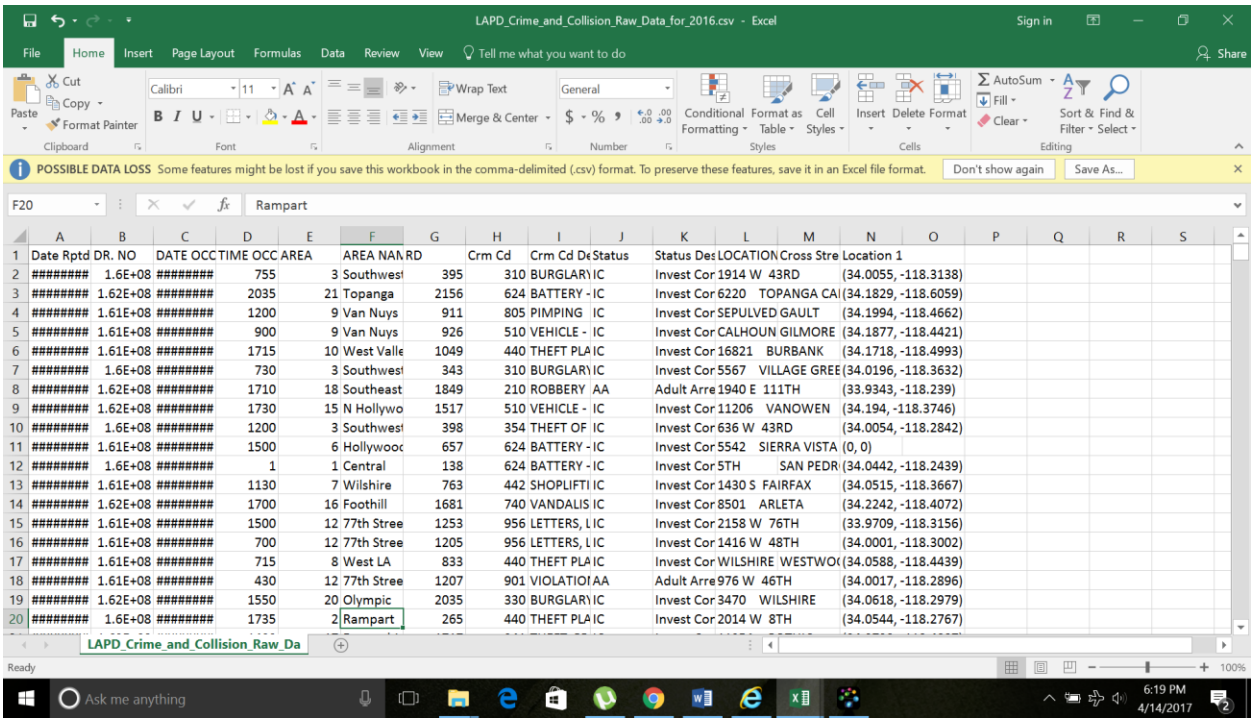
3.Description of data sources

For this project, we are using two large datasets

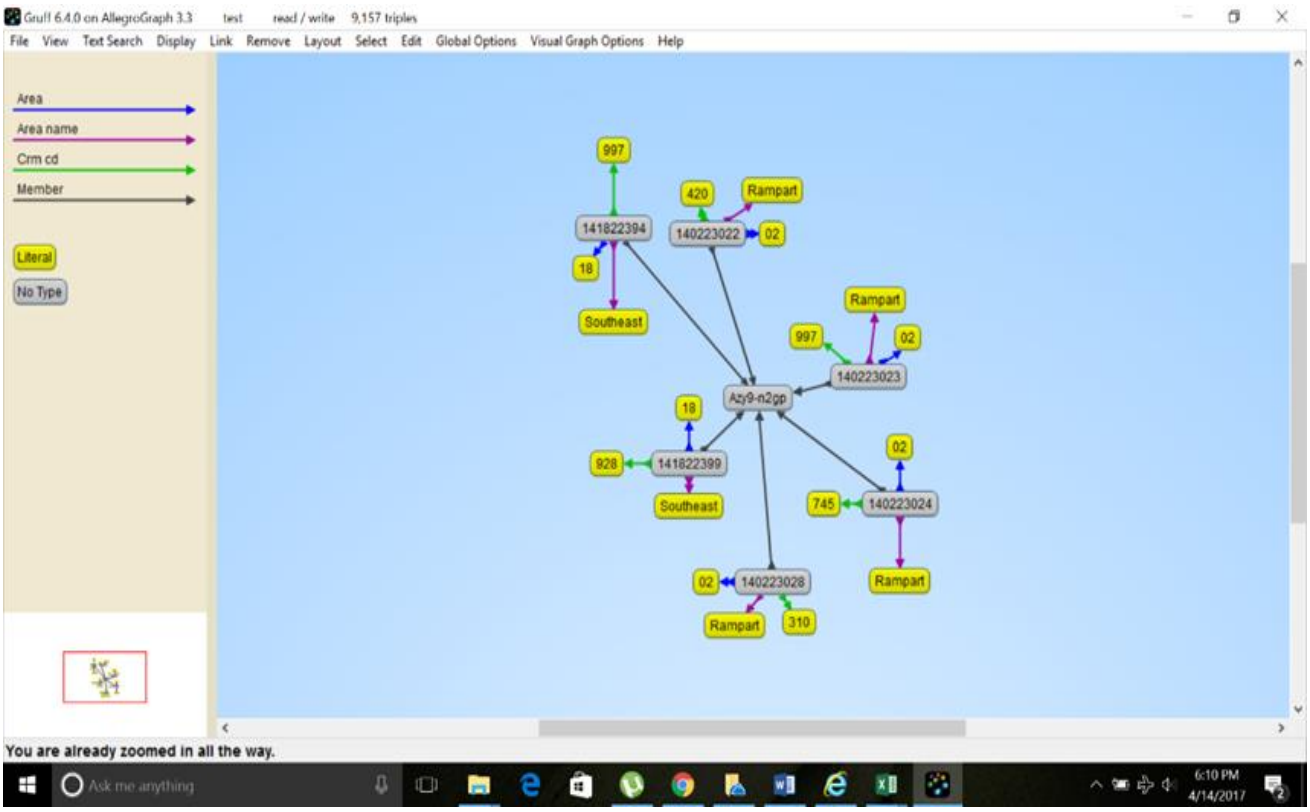
1. LAPD Crime and Collision Data in 2014 which has the following attributes:
    a) Date of crime reporting
    b) Date Occurrence
    c) Time Occurrence
    d) Area
    e) Crime Code Description
    f) Status
    g) Status Description
    h) Location
    i) Cross Street
    j) Location
2. LAPD Crime and Collision Data in 2014 which has the following attributes:
    a) Date of crime reporting
    b) Date Occurrence
    c) Time occurrence
    d) Area
    e) Crime Code Description
    f) Status
    g) Status Description
    h) Location
    i) Cross Street
    j) Location



Crime dataset 2014 (above)



Crime dataset 2016 (above)



Ontology information

4.Data Integration

The two datasets Crime 2014 and Crime 2016 are hosted in a local Fuseki Server where we look upon the query patterns. The graphs of the different datasets are properly analyzed to present interesting inferences. The query patters are extracted and saved locally which is then used by our web page to visualize interesting results. The visualization part happens simultaneously and there is no lag between the Fuseki server and the web page which is hosted.

We have included the query we use below.

```
Query :
PREFIX g1:<http://crime2016/>
PREFIX g2:<http://crime2014/>
PREFIX p2:<http://data.lacity.org/resource/azy9-n2gp/>
PREFIX p1:<http://data.lacity.org/resource/ttiz-7an8/>
SELECT ?o (COUNT(?s1) as ?g1count) (COUNT(?s2) as ?g2count) WHERE
{
{GRAPH g1:{?s1 p1:status_desc ?o . }} UNION {
GRAPH g2:{?s2 p2:status_desc ?o . }}
FILTER regex(?o,"ARREST","i")
}
GROUP BY ?o

2.Area where crime has decreased
PREFIX g1:<http://crime2016/>
```

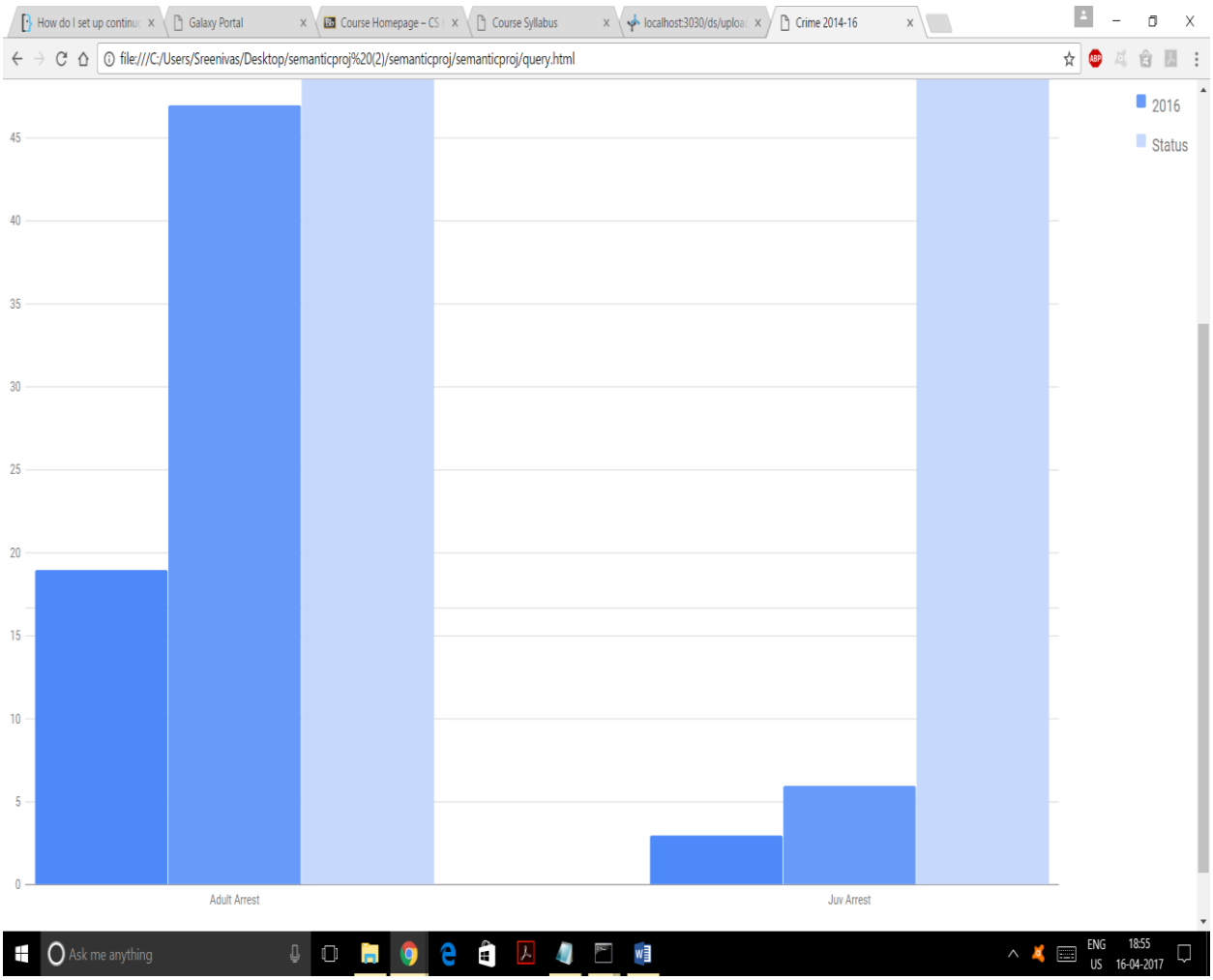
```
PREFIX g2:<http://crime2014/>
PREFIX p2:<http://data.lacity.org/resource/azy9-n2gp/>
PREFIX p1:<http://data.lacity.org/resource/ttiz-7an8/>
SELECT ?o (COUNT(?s1) as ?g1count) (COUNT(?s2) as ?g2count) if((?g1count-
?g2count)>0,"Increased","Decreased") as ?status  WHERE
{
{
GRAPH g1:{?s1 p1:area_name ?o} } UNION {
GRAPH g2:{?s2 p2:area_name ?o} }.
}
GROUP BY ?o

3.Types of crime reporting
PREFIX g1:<http://crime2016/>
PREFIX g2:<http://crime2014/>
PREFIX p2:<http://data.lacity.org/resource/azy9-n2gp/>
PREFIX p1:<http://data.lacity.org/resource/ttiz-7an8/>
SELECT ?o (COUNT(?s1) as ?g1count) (COUNT(?s2) as ?g2count)  WHERE
{
{
GRAPH g1:{?s1 p1:crm_cd_desc ?o} } UNION {
GRAPH g2:{?s2 p2:crm_cd_desc ?o} }.
}
GROUP BY
 ?o
```

5.Data Product Results

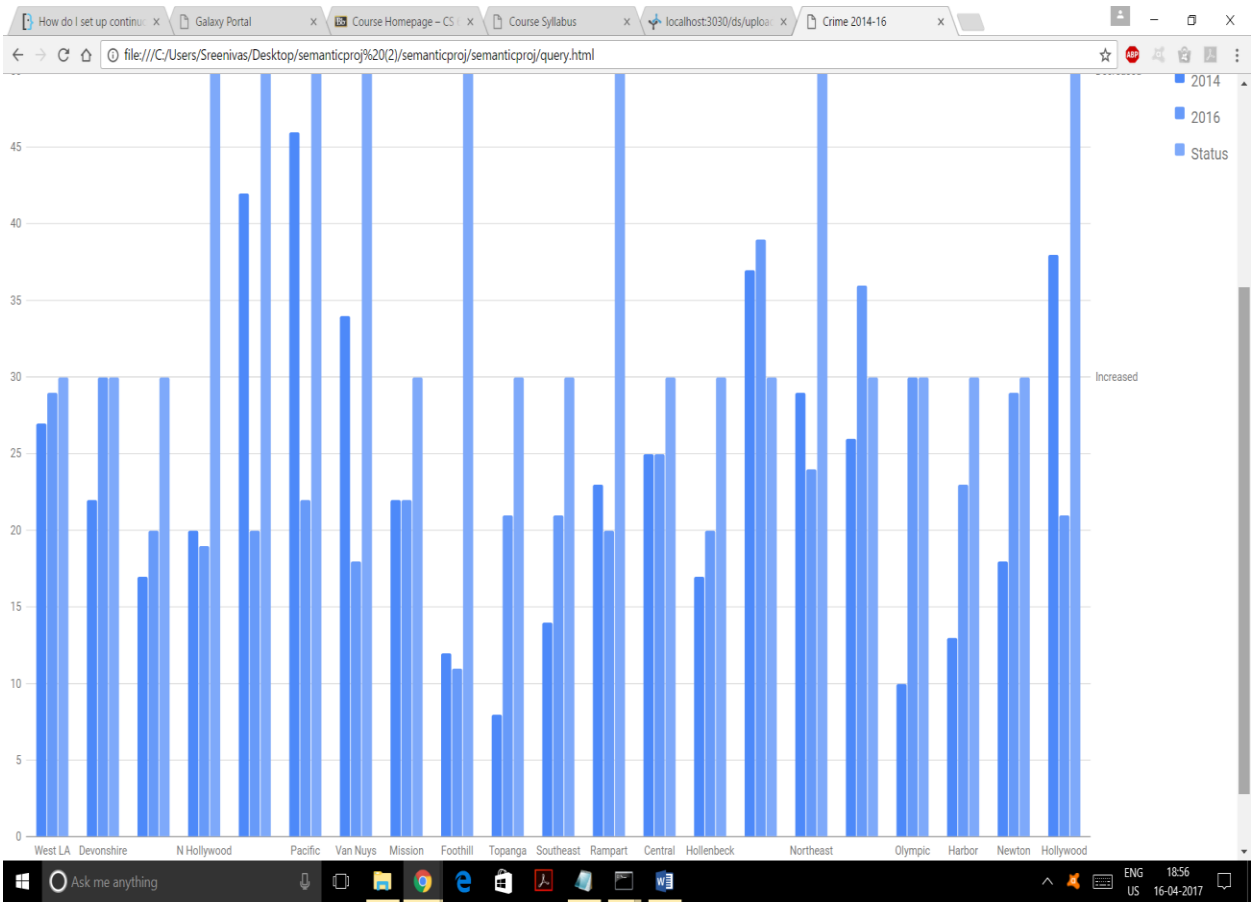
The results are shown in our webpage which is catered to the end users. In our case, the end users are the police department. The different patterns obtained through the dataset are shown in the form of visualizations where the end users can infer the results from. A few of the visualizations are

- Police arrests record
- Crimes per area
- Type of crime reported

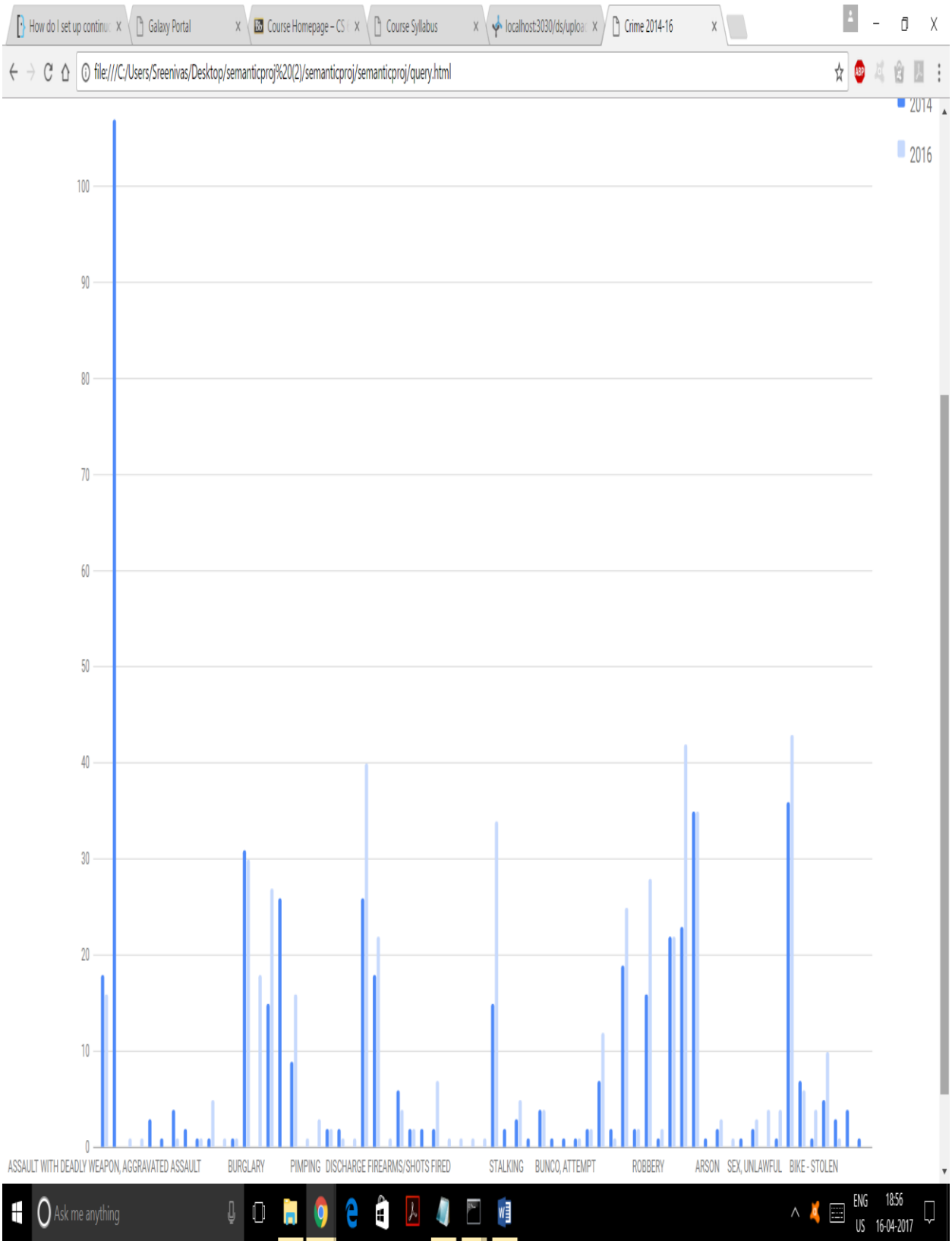
The above three visualizations are shown below.



Police arrests record



Crimes per area



Type of Crime reporting (above)

6.Justification

With a single dataset, we couldn't derive meaningful results. Thus, had to move to a Custom project to integrate multiple datasets. We are using two datasets: Crime rate of Los Angeles for the year 2014 and Crime rate of Los Angeles for the year 2016. We have compared the Crime rates between two years for the city of Los Angeles and proposed the result.

7.Summary

Using Semantic Web technologies like Fuseki for hosting RDF files, HTML and JavaScript for visualization, we have derived useful information from the datasets. We also got the information on where the safest places are, which crimes are occurring often. We have visualized the same through our webpage.