

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

LAPD_Crime_and_Collision_Raw_Data_for_2014.csv - Excel

File Home Insert Page Layout Formulas Data Review View Tell me what you want to do

Clipboard Font Alignment Number Styles Cells Editing

POSSIBLE DATA LOSS Some features might be lost if you save this workbook in the comma-delimited (.csv) format. To preserve these features, save it in an Excel file format. Don't show again Save As...

N2 (34.0429, -118.2499)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
	Date Rptd	DR NO	DATE OCC	TIME OCC	AREA	AREA NAME	RD	Crm Cd	Crm Cd Desc	Status	Des LOCATION	Cross Street	Location 1					
1	#####	1.4E+08	#####	1855	1	Central	176	210	ROBBERY	IC	Invest Cor	7TH MAPLE	(34.0429, -118.2499)					
2	#####	1.4E+08	#####	605	1	Central	156	230	ASSAULT VAO	Adult Oth	Invest Cor	6TH SAN JULI	(34.0432, -118.2466)					
3	#####	1.4E+08	#####	1855	1	Central	176	210	ROBBERY	IC	Invest Cor	7TH MAPLE	(34.0429, -118.2499)					
4	#####	1.4E+08	#####	1200	1	Central	157	649	DOCUMENTIC	Invest Cor	500 S SAN PEDRO	(34.0442, -118.2439)						
5	#####	1.4E+08	#####	1630	1	Central	161	442	SHOPLIFTI	IC	Invest Cor	700 S FIGUEROA	(34.0494, -118.2599)					
6	#####	1.4E+08	#####	2150	1	Central	142	997	TRAFFIC D	IC	Invest Cor	HO GRAND	(34.0519, -118.2541)					
7	#####	1.4E+08	#####	1910	1	Central	192	997	TRAFFIC D	IC	Invest Cor	HO VENICE	(34.0341, -118.2692)					
8	#####	1.4E+08	#####	1400	1	Central	143	442	SHOPLIFTI	IC	Invest Cor	400 S BROADWA	(34.0495, -118.2494)					
9	#####	1.4E+08	#####	1955	1	Central	111	997	TRAFFIC D	IC	Invest Cor	MA ARCADIA	(34.0557, -118.2402)					
10	#####	1.4E+08	#####	1200	1	Central	176	350	THEFT, PEI	IC	Invest Cor	7TH WALL	(34.0429, -118.2499)					
11	#####	1.4E+08	#####	900	1	Central	101	341	THEFT-GR	IC	Invest Cor	900 FIGUEROA	(34.0663, -118.2441)					
12	#####	1.4E+08	#####	15	1	Central	123	624	BATTERY -	IC	Invest Cor	200 W TEMPLE	(34.0549, -118.2426)					
13	#####	1.4E+08	#####	540	1	Central	101	330	BURGLARY	IC	Invest Cor	W MARVIEW	(34.0672, -118.2486)					
14	#####	1.4E+08	#####	1530	1	Central	176	745	VANDALIS	IC	Invest Cor	9TH SAN PED	(34.0404, -118.2537)					
15	#####	1.4E+08	#####	800	1	Central	192	230	ASSAULT V	IC	Invest Cor	181 OLIVE	(34.0358, -118.2708)					
16	#####	1.4E+08	#####	1130	1	Central	174	997	TRAFFIC D	IC	Invest Cor	S LOE 8TH	(34.0413, -118.2532)					
17	#####	1.4E+08	#####	1100	1	Central	156	997	TRAFFIC D	IC	Invest Cor	SAI 5TH	(34.0437, -118.2456)					
18	#####	1.4E+08	#####	1500	1	Central	161	997	TRAFFIC D	IC	Invest Cor	9TH HARBOR	(34.0477, -118.266)					

Ready

Crime dataset 2014 (above)

LAPD_Crime_and_Collision_Raw_Data_for_2016.csv - Excel

File Home Insert Page Layout Formulas Data Review View Tell me what you want to do

Clipboard Font Alignment Number Styles Cells Editing

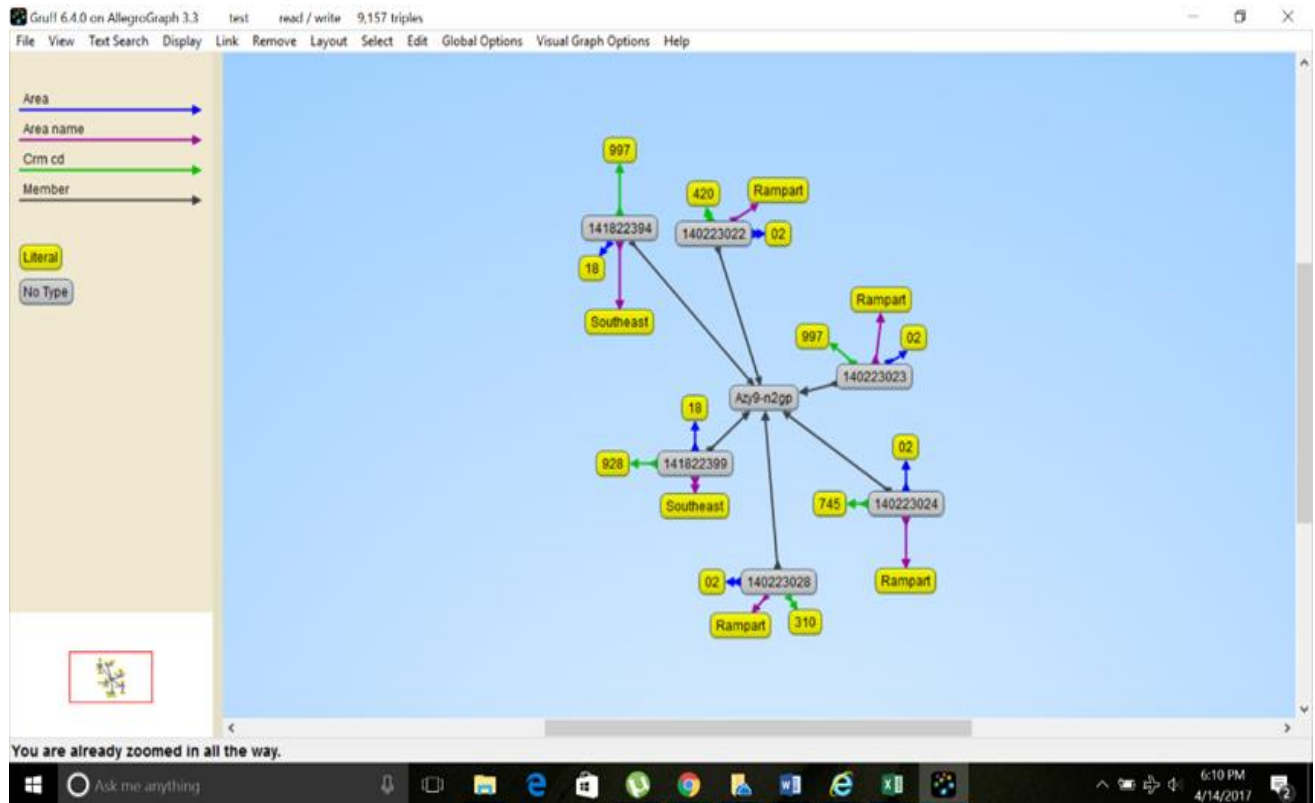
POSSIBLE DATA LOSS Some features might be lost if you save this workbook in the comma-delimited (.csv) format. To preserve these features, save it in an Excel file format. Don't show again Save As...

F20 Rampart

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
	Date Rptd	DR NO	DATE OCC	TIME OCC	AREA	AREA NAME	RD	Crm Cd	Crm Cd Desc	Status	Des LOCATION	Cross Street	Location 1						
1	#####	1.6E+08	#####	755	3	Southwest	395	310	BURGLARY	IC	Invest Cor	1914 W 43RD	(34.0055, -118.3138)						
2	#####	1.62E+08	#####	2035	21	Topanga	2156	624	BATTERY -	IC	Invest Cor	6220 TOPANGA CAI	(34.1829, -118.6059)						
3	#####	1.61E+08	#####	1200	9	Van Nuys	911	805	PIMPING	IC	Invest Cor	SEPULVED GAULT	(34.1994, -118.4662)						
4	#####	1.61E+08	#####	900	9	Van Nuys	926	510	VEHICLE -	IC	Invest Cor	CALHOUN GILMORE	(34.1877, -118.4421)						
5	#####	1.61E+08	#####	1715	10	West Valle	1049	440	THEFT PLA	IC	Invest Cor	16821 BURBANK	(34.1718, -118.4993)						
6	#####	1.6E+08	#####	730	3	Southwest	343	310	BURGLARY	IC	Invest Cor	5567 VILLAGE GREE	(34.0196, -118.3632)						
7	#####	1.62E+08	#####	1710	18	Southeast	1849	210	ROBBERY	AA	Adult Arre	1940 E 111TH	(33.9343, -118.239)						
8	#####	1.62E+08	#####	1730	15	N Hollywo	1517	510	VEHICLE -	IC	Invest Cor	11206 VANOWEN	(34.194, -118.3746)						
9	#####	1.6E+08	#####	1200	3	Southwest	398	354	THEFT OF	IC	Invest Cor	636 W 43RD	(34.0054, -118.2842)						
10	#####	1.61E+08	#####	1500	6	Hollywooc	657	624	BATTERY -	IC	Invest Cor	5542 SIERRA VISTA	(0, 0)						
11	#####	1.6E+08	#####	1	1	Central	138	624	BATTERY -	IC	Invest Cor	5TH SAN PEDRO	(34.0442, -118.2439)						
12	#####	1.61E+08	#####	1130	7	Wilshire	763	442	SHOPLIFTI	IC	Invest Cor	1430 S FAIRFAX	(34.0515, -118.3667)						
13	#####	1.62E+08	#####	1700	16	Foothill	1681	740	VANDALIS	IC	Invest Cor	8501 ARLETA	(34.2242, -118.4072)						
14	#####	1.61E+08	#####	1500	12	77th Stree	1253	956	LETTERS, L	IC	Invest Cor	2158 W 76TH	(33.9709, -118.3156)						
15	#####	1.61E+08	#####	700	12	77th Stree	1205	956	LETTERS, L	IC	Invest Cor	1416 W 48TH	(34.0001, -118.3002)						
16	#####	1.61E+08	#####	715	8	West LA	833	440	THEFT PLA	IC	Invest Cor	WILSHIRE WESTWOC	(34.0588, -118.4439)						
17	#####	1.61E+08	#####	430	12	77th Stree	1207	901	VIOLATIO	AA	Adult Arre	976 W 46TH	(34.0017, -118.2896)						
18	#####	1.62E+08	#####	1550	20	Olympic	2035	330	BURGLARY	IC	Invest Cor	3470 WILSHIRE	(34.0618, -118.2979)						
19	#####	1.6E+08	#####	1735	2	Rampart	265	440	THEFT PLA	IC	Invest Cor	2014 W 8TH	(34.0544, -118.2767)						

Ready

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 patterns 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

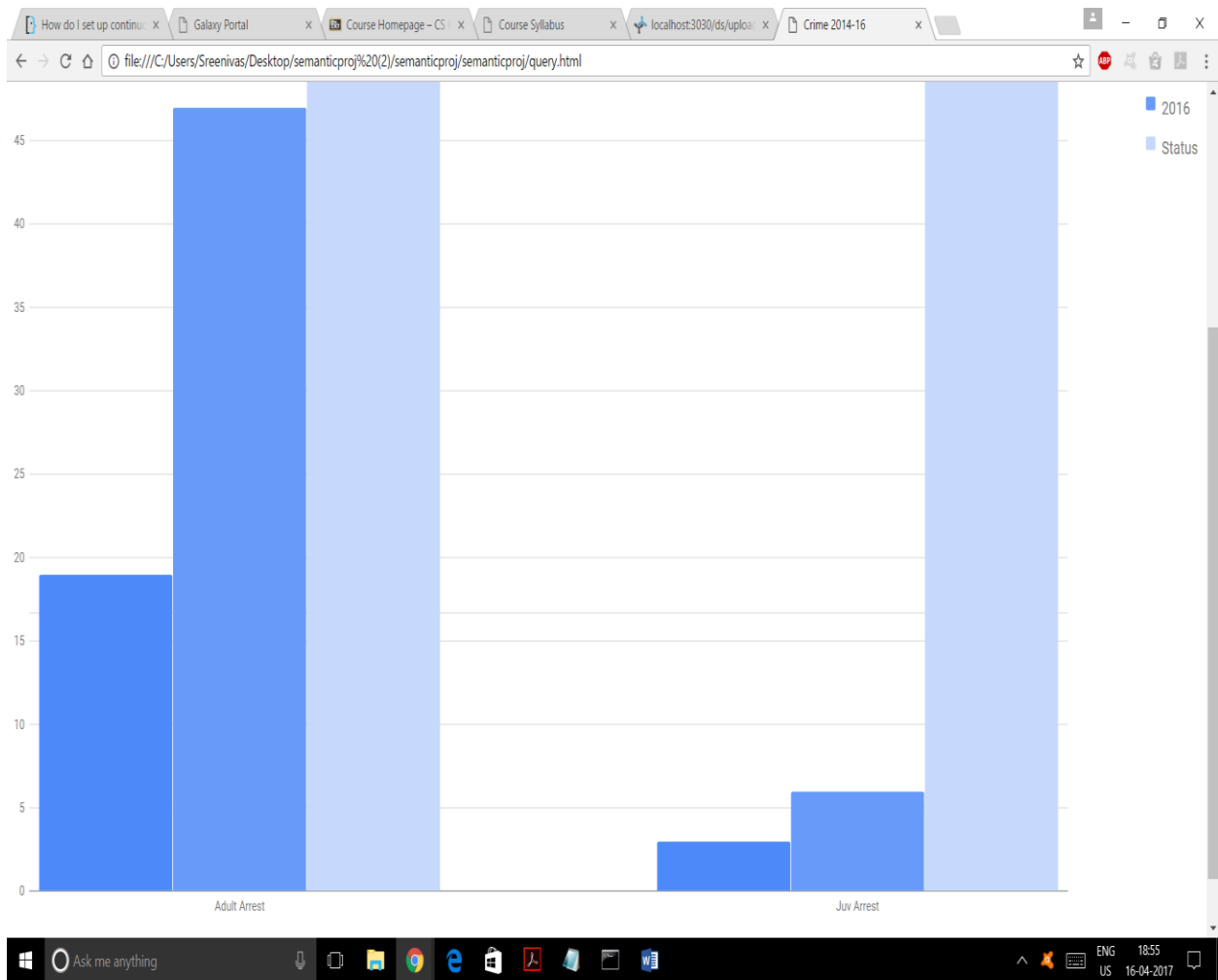
```

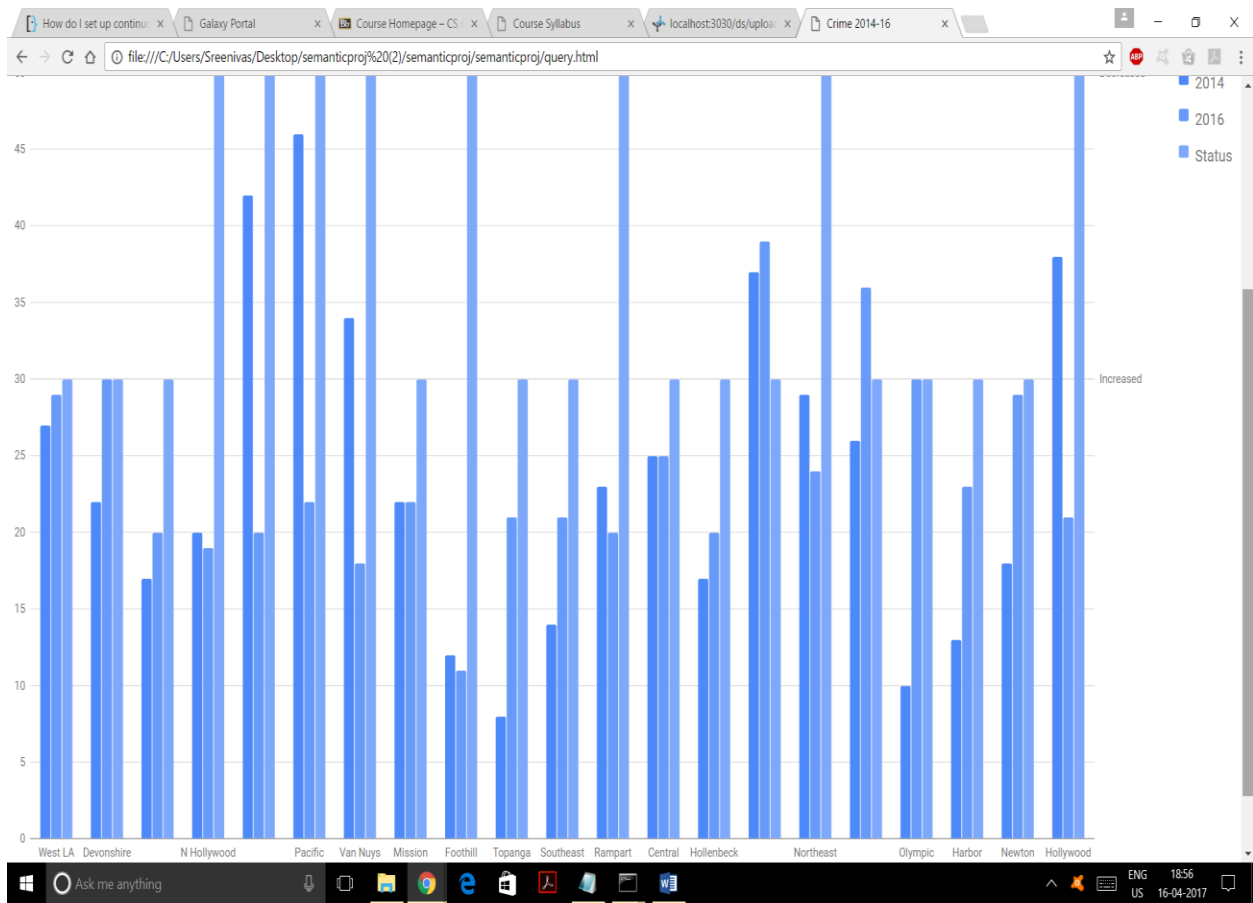
5.Data Product Results

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

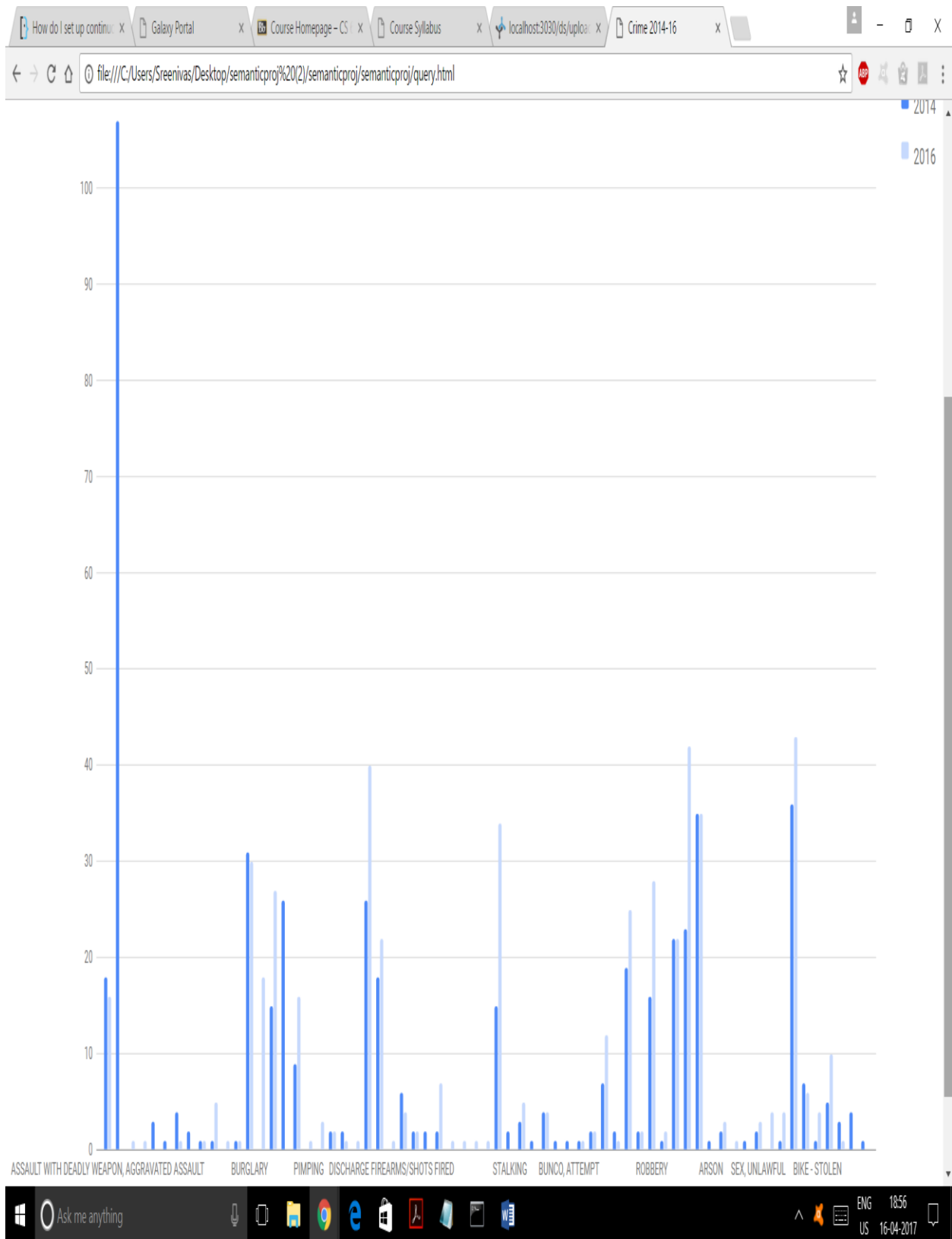
- Police arrests record
- Crimes per area
- Type of crime reported

The above three visualizations are shown below.





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.