Project Report

CS6315 - Semantic Web

Crime Incident Analysis – Seattle Police Department

Team Information:

Team Name: Nocturnals

Team Members:

- Arpita Mothukuri (axm163631)
- Kartheek Anumolu (kxa152730)
- Lakshmi Priyanka Parimi (lxp160730)
- Srinivasa Rajesh Ayachitula (sxa172830)

Type of Project: Custom Project

1. Introduction:

The incidents in Seattle Police Department Police Report Incident dataset are based on initial police reports taken by officers when responding to incidents around the city. The information enters our Records Management System (RMS) and is then transmitted out to data.seattle.gov. This information is published within 6 to 12 hours after the report is filed into the system.

The other dataset in consideration is all the Police responses to 9-1-1 calls within the city. Police response data shows all officers dispatched. To protect the security of a scene, the safety of officers and the public, and sensitive ongoing investigation, these events are added to the data.seattle.gov only after the incident is considered safe to close out. Data is refreshed on a 4 hour interval.

2. Target Audience:

Our analysis on the Crime Incidents will we be helpful to the Seattle Police department. "Seattle Police Department 911 Incident Response" and "Seattle Police Department Police Report Incident" datasets have been joined on incident type. It will give them a view on how many incidents have been reported and on how many actions has been taken. This is very useful for the Seattle police department to analyze their efficiency. This analysis also provides an overview to people living in Seattle about the efficiency of their local police Department. In addition to this we have also provided analysis which shows the type of crime incidents that have occurred at a point of place.

Sector wise analysis on number of incidents reported and responded is displayed on pie chart. Pie chart also displays percentage of incidents reported and responded for each sector.This analysis will also be helpful for the detectives and other Anti-Crime organizations so that they can analyze the number of incidents reported in a year and type of incidents reported.

3. Description of Data Sources

Datasets under consideration are:

1) Seattle Police Department 911 Incident Response (No. of triples: 10025)

Data is available in below location-

https://catalog.data.gov/dataset/seattle-police-department-911-incident-response-52779

 2) Seattle Police Department Police Report Incident (No. of triples: 12227)

Data is available in below location-

https://catalog.data.gov/dataset/seattle-police-department-police-report-incident-7430a

Seattle Police Department 911 Incident Response dataset:

<rdf:Description rdf:about="http://data.seattle.gov/resource/_3k2p-39jp/15736">

   <socrata:rowID>1</socrata:rowID>

   <rdfs:member rdf:resource="https://data.seattle.gov/resource/_3k2p-39jp"/>

   <ds:cad_cdw_id>15736</ds:cad_cdw_id>

   <ds:cad_event_number>10000246357</ds:cad_event_number>

   <ds:general_offense_number>2010246357</ds:general_offense_number>

   <ds:event_clearance_code>242</ds:event_clearance_code>

   <ds:event_clearance_description>FIGHT DISTURBANCE</ds:event_clearance_description>

   <ds:event_clearance_subgroup>DISTURBANCES</ds:event_clearance_subgroup>

   <ds:event_clearance_group>DISTURBANCES</ds:event_clearance_group>

   <ds:event_clearance_date>2010-07-17T20:49:00</ds:event_clearance_date>

   <ds:hundred_block_location>3XX BLOCK OF PINE ST</ds:hundred_block_location>

   <ds:district_sector>M</ds:district_sector>
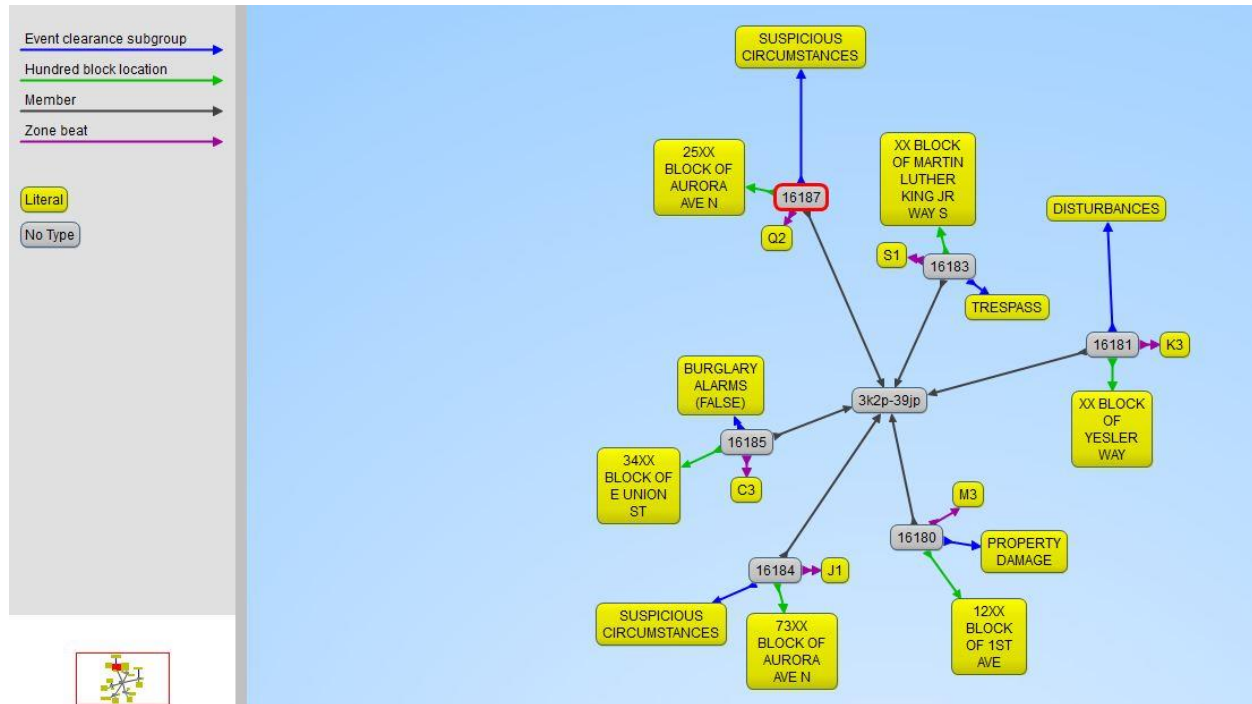
   <ds:zone_beat>M2</ds:zone_beat>

   <ds:census_tract>8100.2001</ds:census_tract>

   <ds:longitude>-122.338146748</ds:longitude>

   <ds:latitude>47.610975163</ds:latitude>

```
<ds:incident_location><geo:SpatialThing>
   <geo:lat>47.610975163</geo:lat>
   <geo:long>-122.338146748</geo:long></geo:SpatialThing></ds:incident_location>
</rdf:Description>
```


Fig1. Displaying some sample triple from Incident Response dataset

Seattle Police Department Police Report Incident:

```
<dsbase:_7ais-f98f rdf:about="http://data.seattle.gov/resource/policereport/887754">
   <socrata:rowID>1401466</socrata:rowID>
   <rdfs:member rdf:resource="https://data.seattle.gov/resource/policereport"/>
   <ds:rms_cdw_id>887754</ds:rms_cdw_id>
   <ds:general_offense_number>201378203</ds:general_offense_number>
   <ds:offense_code>2299</ds:offense_code>
   <ds:offense_code_extension>2</ds:offense_code_extension>
   <ds:offense_type>BURGLARY-SECURE PARKING-RES</ds:offense_type>
   <ds:summary_offense_code>2299</ds:summary_offense_code>
   <ds:summarized_offense_description>BURGLARY-SECUREPARKING-
RES</ds:summarized_offense_description>
```

```xml
<ds:date_reported>2013-03-07T16:29:00</ds:date_reported>
<ds:occurred_date_or_date_range_start>2013-03-
07T16:29:00</ds:occurred_date_or_date_range_start>
<ds:hundred_block_location>60XXBLOCK AV N</ds:hundred_block_location>
<ds:district_sector>B</ds:district_sector>
<ds:zone_beat>B2</ds:zone_beat>
<ds:census_tract_2000>3500.2037</ds:census_tract_2000>
<ds:longitude>-122.354377747</ds:longitude>
<ds:latitude>47.672828674</ds:latitude>
<ds:location><geo:SpatialThing>
    <geo:lat>47.672828674</geo:lat>
    <geo:long>-122.354377747</geo:long></geo:SpatialThing></ds:location>
<ds:month>3</ds:month>
<ds:year>2013</ds:year></dsbase:_7ais-f98f>
```


Fig2. Displaying some sample triples from dataset Incident Reports

4. Data Integration

Two datasets are integrated to give us some critical data. The combined datasets give us enough information to perform an analysis on incidents in the Seattle area.

Apache JENA Fuseki: SPARQL server. Provides REST-style SPARQL HTTP Update, SPARQL Query, and SPARQL Update using the SPARQL protocol over HTTP.
Combining Data:

- Datasets Incident Report and Incident Response are integrated to retrieve various results like sectors with most incident reports, sectors with most incident responses, years with most incident reports and the types of incidents being reported and resolved.

Table 1: Attributes used for combining data

Attribute	Description
/summarized_offense_description	Summary of the offense during incident report
/latitude	Latitude of the incident reported
/longitude	Longitude of the incident reported
/year	Year when the incident is reported
/district_sector	District sector of incidents reported
/_3k2p-39jp/district_sector	District sector of incidents responded to
/event_clearance_subgroup	Summary of incident clearance
/offense_type	Type of incident offense committed

5. Data Product Results

SPARQL Queries:

The data is being integrated using the "**state**" property in every dataset. The SPARQL queries shown below are executed in FUSEKI server's SPARQL endpoint and the corresponding results are retrieved:

Query 1: This query is used to retrive the type and location of the incidents reported in the Seattle area.

SELECT ?latitude ?longitude ?offense WHERE{ ?subject
<https://data.seattle.gov/resource/policereport/summarized_offense_description> ?offense;
<https://data.seattle.gov/resource/policereport/latitude> ?latitude;
<https://data.seattle.gov/resource/policereport/longitude> ?longitude; }

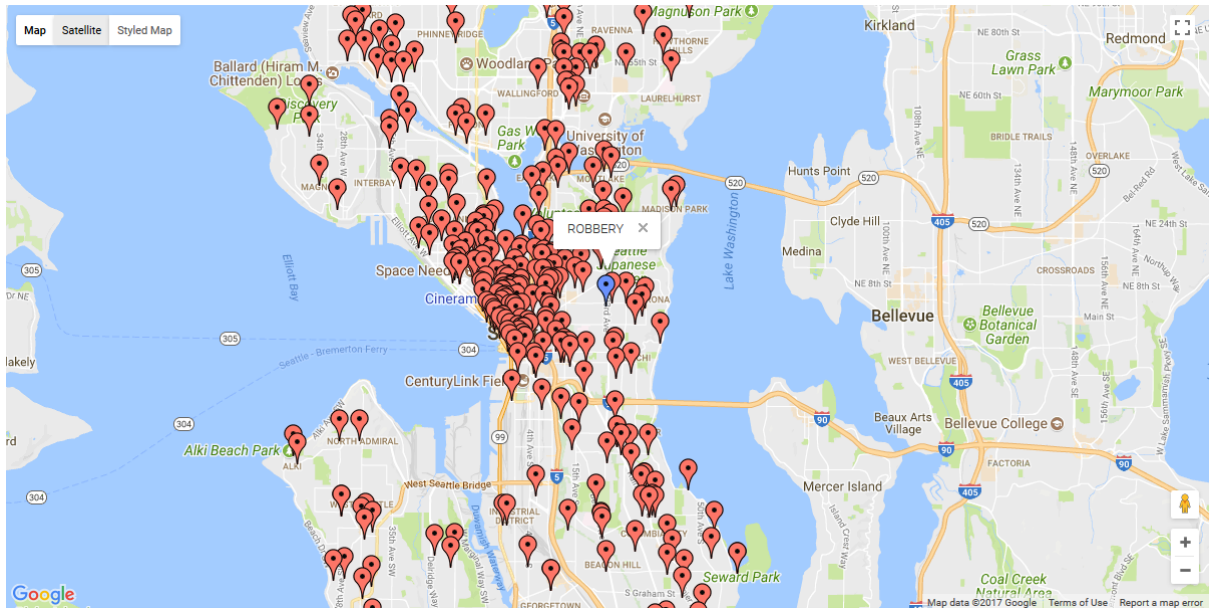**Fig 3.** Screenshot of results for Query 1 in Fuseki Server

Fig 4. Visualization results for Query 1

Query2: This query retrives the incidents responded by the 911 which are grouped year wise.

SELECT ?year (count(distinct ?s) as ?responsecount) WHERE {   ?s
<https://data.seattle.gov/resource/policereport/year> ?year.} GROUP BY ?year

	year	responsecount
1	"2016"	"297"^^xsd:integer
2	"2015"	"5"^^xsd:integer
3	"2014"	"9"^^xsd:integer
4	"2013"	"187"^^xsd:integer
5	"2012"	"2"^^xsd:integer

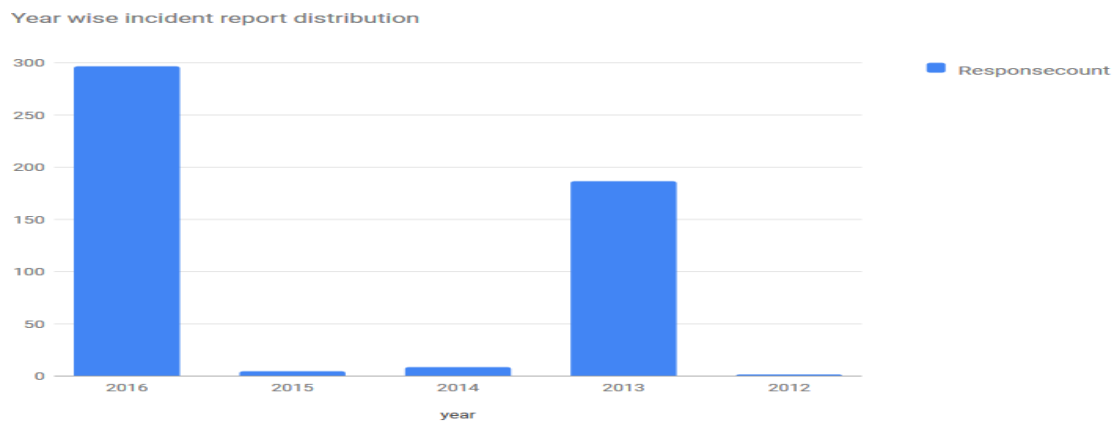**Fig 5.** Screenshot of results for Query 2 in Fuseki Server

Fig 6. Visualization for Query 2

Query 3: This query is used to retrieve incident reports in a sector.

SELECT ?sector (COUNT(?subject) AS ?count) WHERE { ?subject <https://data.seattle.gov/resource/policereport/district_sector> ?sector; } Group BY ?sector

	sector	count
	Showing 1 to 18 of 18 entries	
1	"K"	"35"^^xsd:integer
2	"L"	"29"^^xsd:integer
3	"J"	"17"^^xsd:integer
4	"O"	"11"^^xsd:integer
5	"M"	"47"^^xsd:integer
6	"N"	"27"^^xsd:integer
7	"C"	"33"^^xsd:integer
8	"D"	"34"^^xsd:integer
9	"B"	"23"^^xsd:integer

Fig 7. Screenshot of results for Query 3 in Fuseki Server

Query 4: This query is used to retrieve responded incidents in a given sector.

SELECT ?sector (COUNT(?subject) AS ?count) WHERE { ?subject <https://data.seattle.gov/resource/_3k2p-39jp/district_sector> ?sector; } Group BY ?sector

	sector	count
	Showing 1 to 17 of 17 entries	
1	"K"	"43"^^xsd:integer
2	"L"	"25"^^xsd:integer
3	"J"	"33"^^xsd:integer
4	"O"	"11"^^xsd:integer
5	"M"	"58"^^xsd:integer
6	"N"	"19"^^xsd:integer
7	"C"	"23"^^xsd:integer
8	"D"	"59"^^xsd:integer
9	"B"	"30"^^xsd:integer

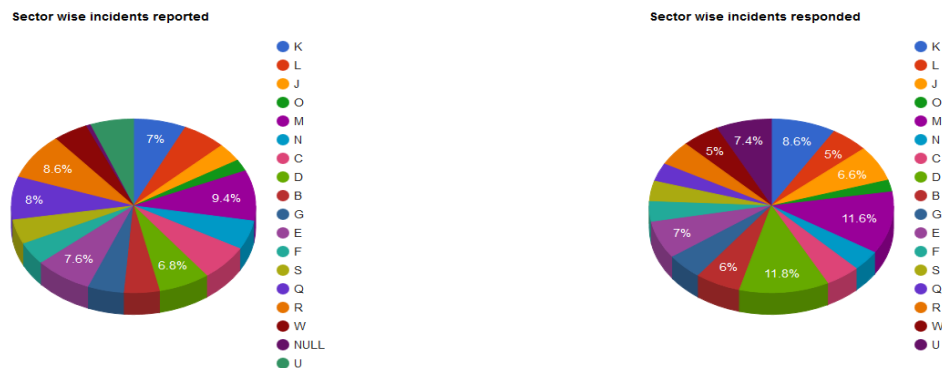**Fig 8**. Screenshot of results for Query 4 in Fuseki Server

Fig 9. Sector incidents reported and responded

Query 5: This query gives us the number of incidents reported and responded in different sectors.

SELECT ?district (count(distinct ?s) as ?responsecount) (count(distinct ?s1) as ?reportcount)WHERE {?s <https://data.seattle.gov/resource/policereport/district_sector> ?district.?s1 <https://data.seattle.gov/resource/_3k2p-39jp/district_sector> ?district}GROUP BY ?district

	district	responsecount	reportcount
	Showing 1 to 17 of 17 entries		Search:
1	"B"	"23"^^xsd:integer	"30"^^xsd:integer
2	"C"	"33"^^xsd:integer	"23"^^xsd:integer
3	"D"	"34"^^xsd:integer	"59"^^xsd:integer
4	"E"	"38"^^xsd:integer	"35"^^xsd:integer
5	"F"	"21"^^xsd:integer	"19"^^xsd:integer
6	"G"	"24"^^xsd:integer	"24"^^xsd:integer
7	"J"	"17"^^xsd:integer	"33"^^xsd:integer

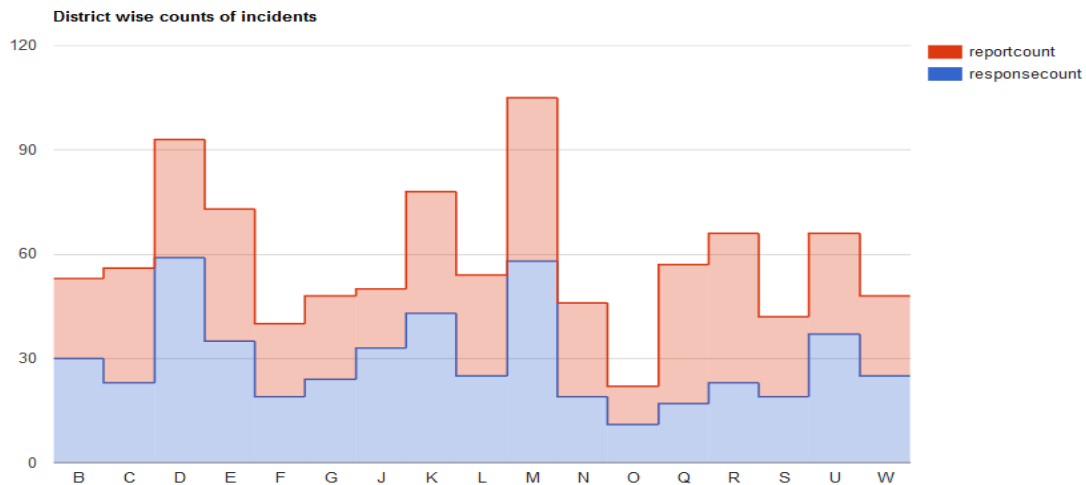**Fig 10.** Screenshot of results for Query 5 in Fuseki Server

Fig 11. District Wise count of incidents

Query 6: This query gives us the number of incidents reported and number of incidents resolved for a given offence type.

SELECT ?offence (COUNT(DISTINCT ?subject1)AS ?reports) (COUNT(DISTINCT ?subject2)AS ?resolved){   ?subject1 <https://data.seattle.gov/resource/policereport/offense_type> ?offence.   ?subject2 <https://data.seattle.gov/resource/_3k2p-39jp/event_clearance_subgroup> ?obj1.   filter(regex(str(?offence), ?obj1)) } Group BY ?offence

	offence	reports	resolved
1	"TRESPASS"	"8"^^xsd:integer	"15"^^xsd:integer
2	"VEH-THEFT-TRUCK"	"3"^^xsd:integer	"16"^^xsd:integer
3	"THEFT-SHOPLIFT"	"26"^^xsd:integer	"16"^^xsd:integer
4	"ROBBERY-BANK-GUN"	"1"^^xsd:integer	"3"^^xsd:integer
5	"FRAUD-IDENTITY THEFT"	"2"^^xsd:integer	"16"^^xsd:integer
6	"ROBBERY-BUSINESS-GUN"	"1"^^xsd:integer	"3"^^xsd:integer
7	"THEFT-LICENSE PLATE"	"3"^^xsd:integer	"16"^^xsd:integer

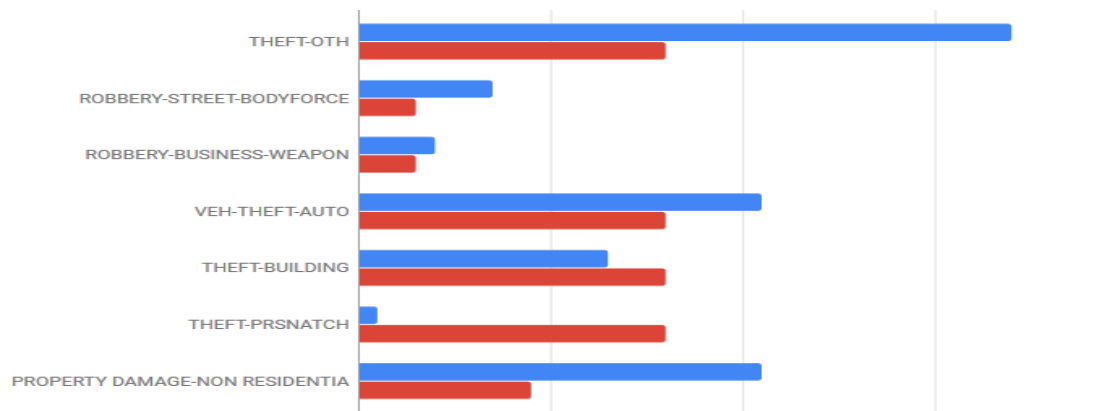**Fig 12.** Screenshot of results for Query 6 in Fuseki Server

Fig 13. Number of incidents reported and resolved for offence type.

6.Custom Project Justification

Crime Incident Analysis can be classified as a custom project with few main reasons. Firstly, Datasets have been uploaded into Apache Jena FUSEKI server and SPARQL end point is created to access these datasets. Apache Jena FUSEKI server uses local host to access the server on web page unlike simple project (i.e logd) where LOGD SPARQL endpoints are directly available. Secondly, the datasets are taken from "data.gov" website and uploaded into FUSEKI server. Whereas in a simple project the datasets are already present in LOGD website. Thirdly, the output from the SPQARL queries is in JSON format which is fed to Google visualization API whereas in simple project query is directly present in LOGD website.

7. Summary

We have started our project by analyzing and reviewing various datasets. Then we finalized on two datasets: "Seattle Police Department 911 Incident Response" and "Seattle Police Department Police Report Incident" from data.gov website. These two datasets have been downloaded (i.e rdf files) and uploaded into Gruff to understand the semantics of these datasets. Then we started looking into integration of these datasets by extracting the attributes for each of the dataset. We then decide to integrate on "Incident type" to provide better analysis on how many incidents are reported and responded.

The finalized datasets (i.e rdf files) have been uploaded into Apache Jena FUSEKI server and SPARQL queries are executed to extract the desired information. The output of these queries is in JSON format which is fed into Google Visualization API. Google Visualization API provides better visualization results. Final web page has been developed using HTML,Bootstrap, Javascript and Google visualization packages Bar, Map, Corechart, Piechart and SteppedAreaChart.