Analysis of Los Angeles Police Department (LAPD) Activity with

Demographic and Economic Implications

Gregory David Mamoyac, Tejas Agara Chandrakumar, Ruchita Shinde, Nikita Shendkar, Zixiao Fu

Department of Information Systems, California State University

Los Angeles

**Abstract:** When faced with the task of analyzing nearly eight million rows of reporting information on LAPD activity and the City of LA’s demographic and economic history, Apache Pig is an effective tool for extracting, transforming, and loading the batch data. Further, Apache Pig allows for the database management and enables for the writing of complex data transformation using high-level MapReduce programming.

The bulk of the data is derived from LAPD activity, which is split into three datasets. The three datasets are vehicle and pedestrian stops, crimes, and arrests information originally collected for purposes of police reporting. The remainder of the data is on LA’s demographic, economic, and census history. The datasets were collected from government open source. Analysis is made along field relations within a single dataset and across multiple datasets in order to discover unique insights that can be applied to other cities with similar data.

The analysis accomplishes unique insights into police activity and city awareness. It continues the on the key themes of public safety, open transparency, accountability, civic engagement, public benefit from government open source, integrity of services, and public resilience. Limitations of analysis are included along with research details and findings.

1. **Introduction**

Analysis is made along field relations within a single dataset and across multiple datasets in order to discover unique insight. This is done with Apache Hadoop IOP 4.2 and Pig version 0.15.0 on a 3- 1 data node- management node configuration. The total storage size of the four datasets is roughly 1.5GB originally collected from government open source at DataLA and ControlPanel LA.

For the purposes of this research, the LAPD activity is specifically broken down into the following: Vehicle and Pedestrian Stop Data 2010 to Present; Crime Data from 2010 to Present; and Arrest Data from 2010 to Present [1]. Further insight will be gained from LA’s Demographic and Economic Statistics [2] .

Analysis is made along field relations within a single dataset and across multiple datasets in order to discover unique insights. The first points of analysis are: which reporting district has the most instances of LAPD activity; which periods of time are most and least affected by LAPD activity; and the relation of crimes that have led to arrests. These findings exhibit insight into police activity broken down into specific fields, allowing for better allocation of police resources in regards to geo-spatial and temporal data.

Further points of analysis are: which demographic profile is the most and least prevalent in LAPD activity; the relation of personal income per capita to LAPD activity; the relation of public school enrollment to LAPD activity; and the relation of unemployment rate to LAPD activity. These findings exhibit insight into city awareness, specifically the correlation of police activity to demographic and economic issues.

While considering certain limitations of analysis, the results of this research can ultimately provide a scope for solutions to police activity. Furthermore, the results can address demographic and economic issues that may directly or at least indirectly correlate to police activity in a given city. This correlation should warrant further study on the causality between a variable factors and the prevalence of policy activity.

1. **Main Results**

While Apache Pig was used for processing analysis on the batch data, Tableau was used for creating the visualizations of our results.

* 1. **Which Reporting District Has the Most Instances of LAPD Activity**

Figure 1. Top Five Reporting Districts Having Instances of Stops, Crimes, Arrests

* 1. **Which Periods of Time are Most and Least Affected by LAPD Activity**
     1. **Which Years are Most and Least Affected by LAPD Activity**

Figure 2. Total Instances of Stops, Crimes, and Arrests for Years 2010-2017

* + 1. **Which Months are Most and Least Affected by LAPD Activity**

Figure 3. Total Instances of Stops, Crimes, and Arrests for Months for Given Year

* 1. **The Relation of Crimes that May Have Led to Arrests**

Figure 4. Comparison of Total Instances of Crimes and Total Instances of Arrests by Years 2010-2017

* 1. **Which Demographic Profile is the Most Prevalent in LAPD Activity**
     1. **Which Descent is Most Affected by LAPD Activity**

Figure 5. Top Five Descents Having Most Instances of Stops, Crimes, and Arrests for Years 2010-2017

* + 1. **Which Age is Most Affected by LAPD Activity**

Figure 6. Top Ten Ages Having Most Instances of Stops, Crimes, and Arrests for Years 2010-2017

* + 1. **Which Sex is Most Affected by LAPD Activity**

Figure 7. Top Sex Having Most Instances of Stops, Crimes, and Arrests for Years 2010-2017

* 1. **The Relation of Personal Income Per Capita to Crimes**

Figure 8. Comparison of Personal Income Per Capita to Crimes for Years 2010-2015

* 1. **The Relation of Public School Enrollment to Crimes**

Figure 9. Comparison of Public School Enrollment to Crimes for Years 2010-2014

* 1. **The Relation of Unemployment Rate to Crimes**

Figure 10. Comparison of Unemployment Rate to Crimes for Years 2010-2017

1. **Conclusion**

**References**

[1] DataLA

[2] ControlPanel LA