Analysis of Los Angeles Police Department (LAPD) Activity with

Demographic and Socio-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 socio-economic history from 2010- present 2017, Apache Pig is an effective tool for extracting, transforming, and loading the data. 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, socio-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 Socio-economic Statistics Data [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 Reports 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 socio-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 address demographic and socio-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 variable sociosocio-economic factors and the prevalence of police activity.

1. **Main Results**

The following section reveal insight on LAPD activity as well as demographic and socio-economic implications related to the such activity. The data encompasses statistical figures in regards to geo- spatial and temporal data.

* 1. **LAPD Reporting Districts with the Most Reports of LAPD Activity**

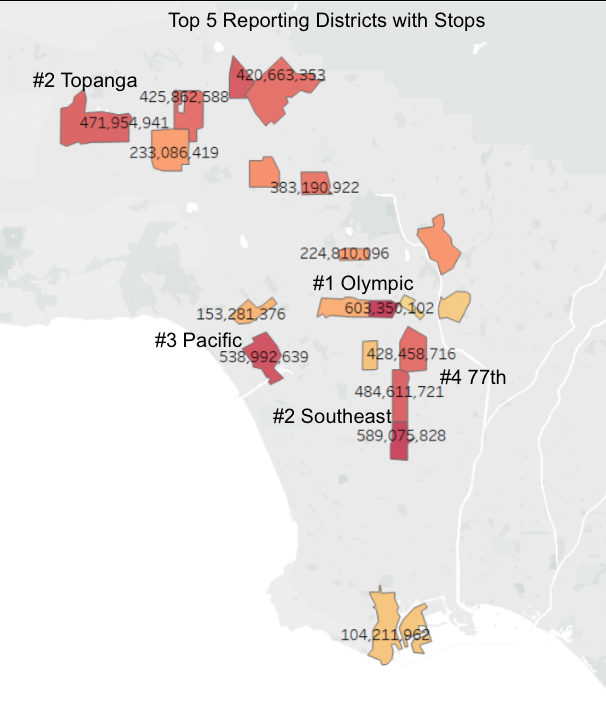
By discovering which LAPD Reporting districts have the most reports of LAPD Activity, a more efficient allocation of resources can be achieved. Therefore, police forces can focus time, manpower, and resources on certain areas. The following data includes data from 2010-Present and lists the top five reporting districts.

Figure 1. Top Five Reporting Districts Having Reports of Stops from 2010-Present

By definition, Stops data encompasses vehicle and pedestrian stops conducted under just cause. Just cause includes traffic violations, crimes, and suspicious activity.

Figure 2. Top Five Reporting Districts Having Reports of Crimes from 2010-Present

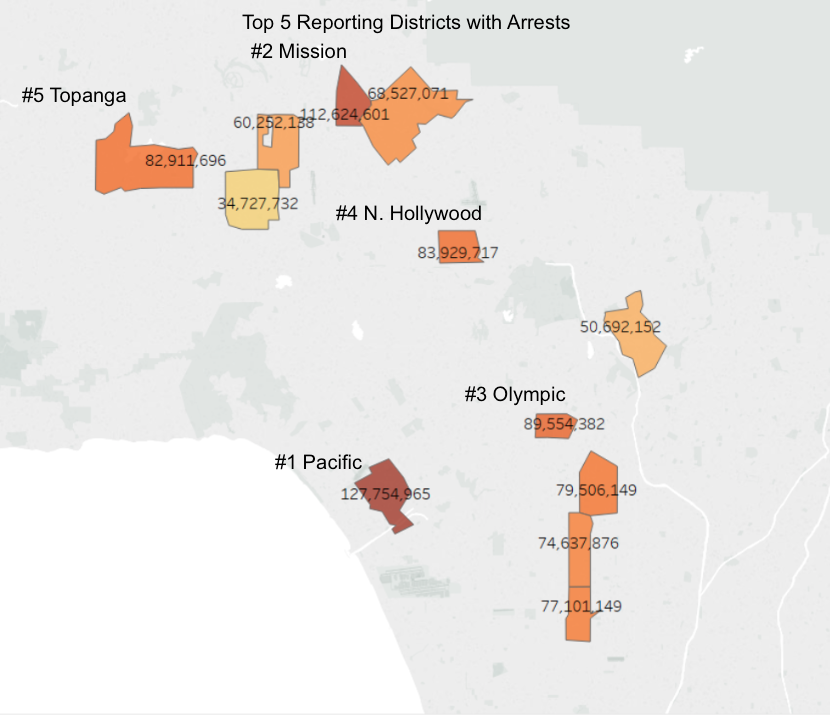
Crimes data encompasses a number law breaking activities such as DUIs, physical assaults, stolen property, and killings. These are reports of crimes and not necessarily first have accounts made by police forces.

Figure 3. Top Five Reporting Districts Having Reports of Arrestsfrom 2010-Present

Arrests define activity where police forces apprehend individuals due to law breaking activity or crimes. Those arrested are then processed into the criminal justice system.

* 1. **Months Most and Least Affected by LAPD Activity**

The following data includes the trends of LAPD activity for months from 2010-Present. This reveals periods in of time when police forces can expect to see high activity. Therefore, resources can be scaled up or down in expectation of such periods.

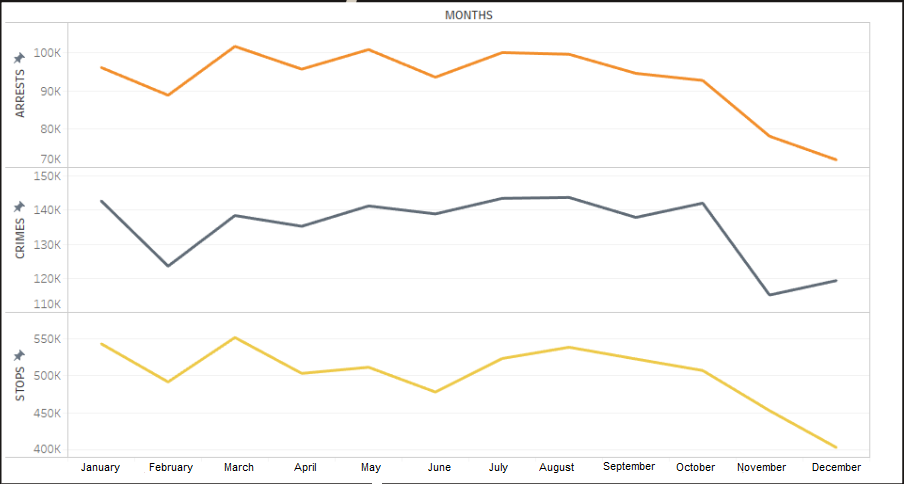
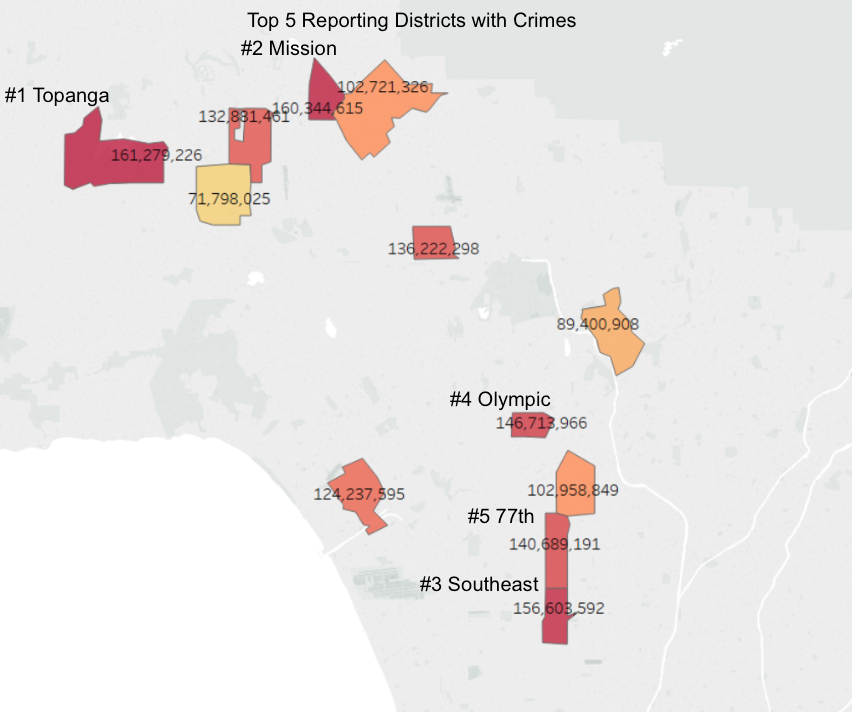


Figure 4. Total Reports of Stops, Crimes, and Arrests Per Month from 2010-Present

The figure encompasses data from Stops, Crimes, and Arrests from 2010-present. The figures reveals dips during February, March to June, and November to December. While there may be several implications of this that potentially include holidays seasons or tax returns, police forces can scale down resources during these periods of time. Inversely, police forces can scale up resources during consistently consistent high activity periods like March.

* 1. **The Relation of Crimes to Arrests**

The relation of the number of reports of crimes to arrests reveals the effectiveness of police forces from 2010-Present. This data reflects police performance in terms of maintaining safe communities and enforcing laws by arresting individuals who have committed crimes.

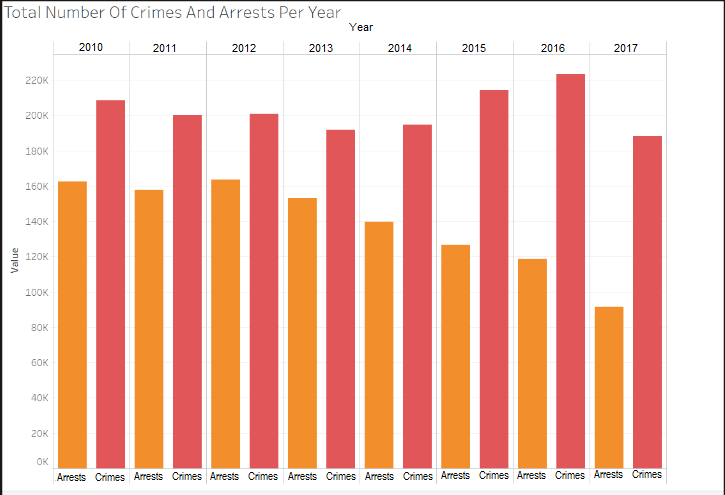


Figure 5. Comparison of Total Reports of Crimes and Total Reports of Arrests by Years 2010-2017

While crimes have remained generally consistent from 2010-Present, the data reveals that arrests have decreased significantly from 2014 - Present. This highlights public interest of community safety and prompts potential action from policymakers and police administrators, and community leaders.

* 1. **Demographic Profile that are Most Prevalent in LAPD Activity Reporting**

Reveals insight into demographic profiles that are most prevalent in LAPD reports. The following figures reveals data for community leaders to rightly alleviate issues that are related to these groups of people in order to pursue accountability and public resilience. Demographic profiles include descent, age, and sex.

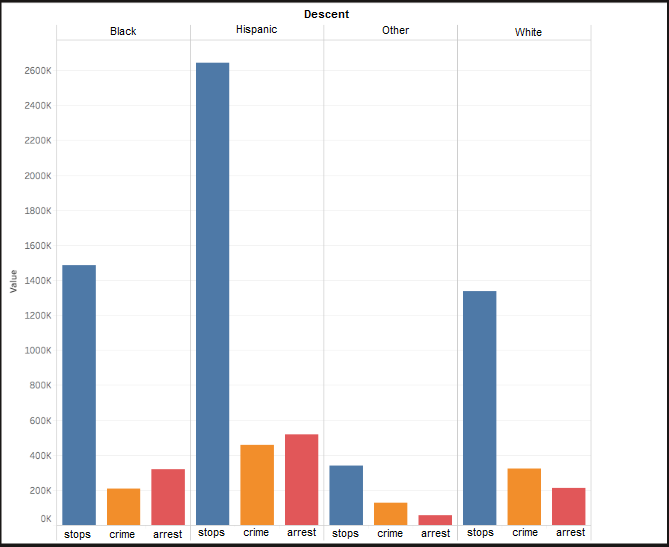


Figure 6. Top Four Descents Having Most Reports of Stops, Crimes, and Arrests for Years 2010-Present

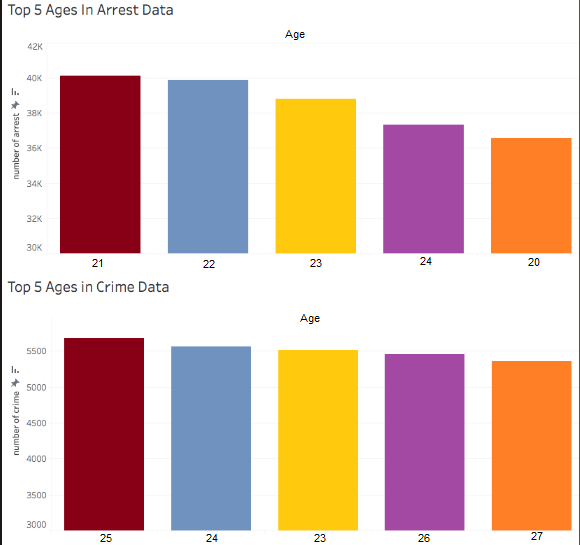


Figure 7. Top Five Ages Having Most Reports Crimes and Arrests for Years 2010-Present

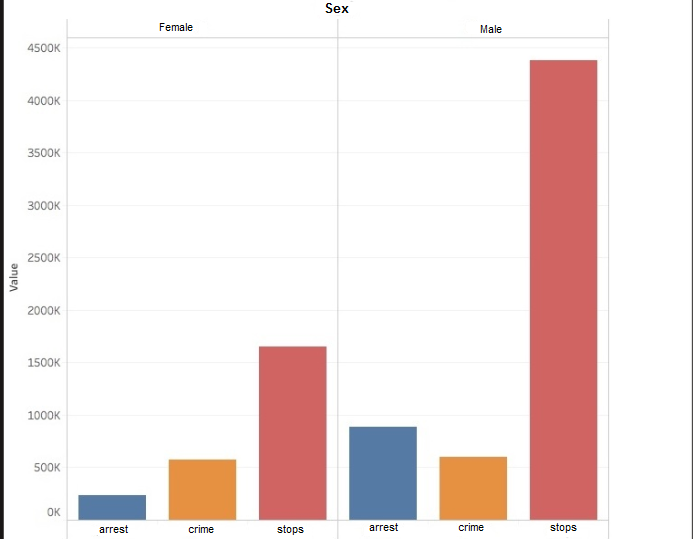


Figure 8. Top Sex Having Most Reports of Stops, Crimes, and Arrests for Years 2010-2017

Examining the top instances of demographic profile in LAPD activity in figures 6-8, the data reveals that Hispanic males in their twenties are most prevalent in reporting data. This data draws attention to community leaders in order to help at-risk youth and improve public welfare.

**2.5 Socio-economic Implications of Crimes**

Socio-economic implications of crimes deals with the factors that potentially cause such activity that goes into LAPD reports. Socio-economic measures such as personal income per capita, public school enrollment, and unemployment are correlated or compared to crimes in relation to years provided by open source data.



Figure 9. Comparison of Personal Income Per Capita to Crimes for Years 2010-2015

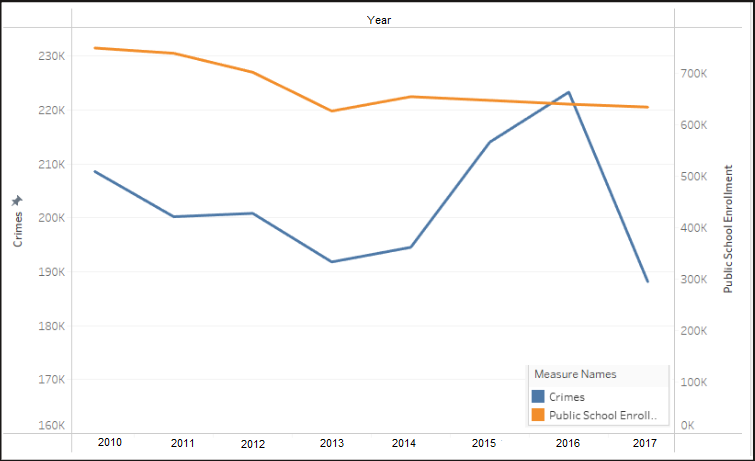
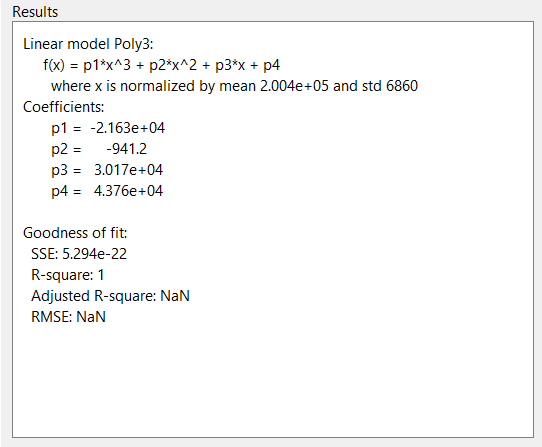


Figure 10. Comparison of Public School Enrollment to Crimes for Years 2010-2014

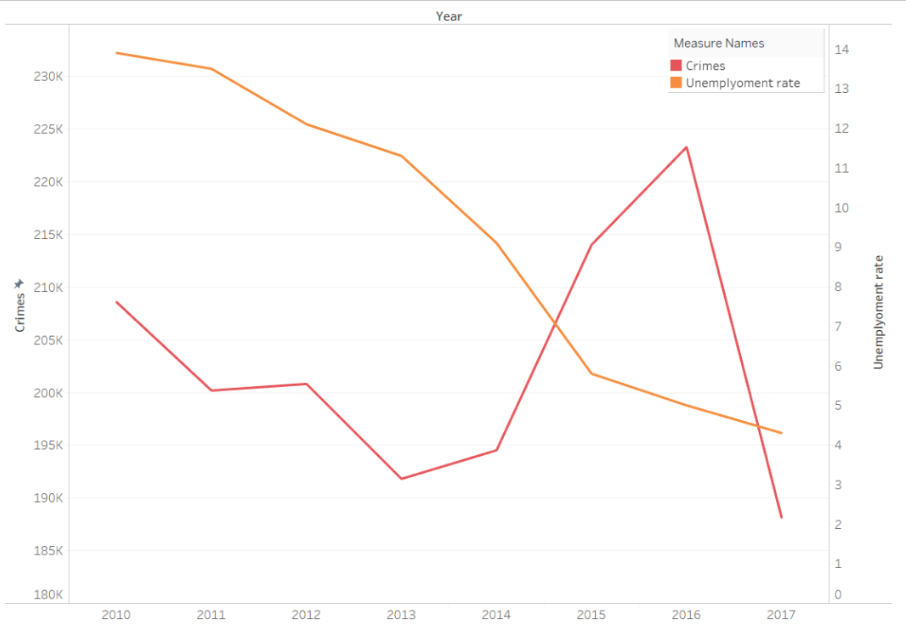


Figure 11. Comparison of Unemployment Rate to Crimes for Years 2010-2017

The above figures 9-11 attempt to reveal socio-economic causes of crime. However, the socio-economic data obtained from open source lacks information such as detailed spatial, month-to-month, and demographic value. This reduces effectiveness of causality implications which are important for policymakers, community leaders, and police department for designing programs for areas of interest. If such values existed, more accurately causality implications would exist to explain relations and thereby address issues and challenges.

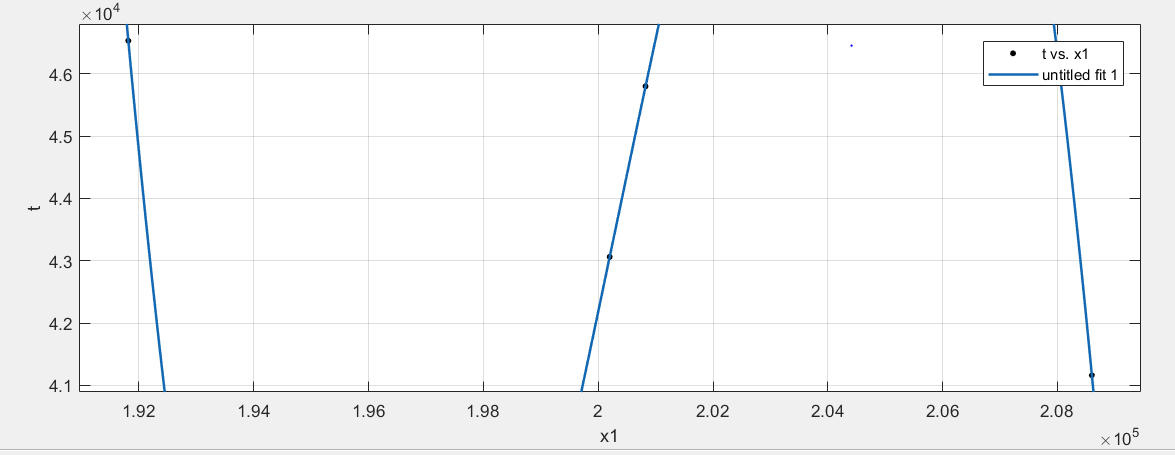


Figure 12. Regression Model for Relation of Crimes to Personal Income Per Capita using Matlab

Figure 13. Poly Model using Matlab

The above figures 12-13 demonstrate the attempted regression model for crimes to personal income per capita. However, the data did not sufficiently satisfy curve fitting function of Matlab to prove an accurate result.

1. **Conclusion**

Analysis into "LAPD Activity” provides insight into significant periods of times and areas of interest where police can focus valuable resources including budgeting, manpower, and energy.

Analysis into "Demographic and Socio-Economic Implications" provides the examination for political and socioeconomic initiatives that may be designed by policymakers and leaders in the community.

This data project continues the key themes of public safety, open transparency, accountability, civic engagement, public benefit from government open source, integrity of services, and public resilience. Furthermore is opens the study of public safety activity that can be applied to other police databases in order to address at-risk groups and communities.

**References**

[1] “Los Angeles - Open Data Portal,” *City of Los Angeles*. [Online]. Available: https://data.lacity.org/browse. [Accessed: 06-Nov-2017].

[2] “Los Angeles | Open Data | Control Panel LA,” *Socrata*. [Online]. Available: https://controllerdata.lacity.org/data. [Accessed: 06-Nov-2017].

[3] “Federal Reserve Economic Data | FRED | St. Louis Fed,” *FRED*. [Online]. Available: https://fred.stlouisfed.org/. [Accessed: 06-Nov-2017].