Descriptive Analysis

**Part A: Daily violence analysis, four weeks before the election, 2013**

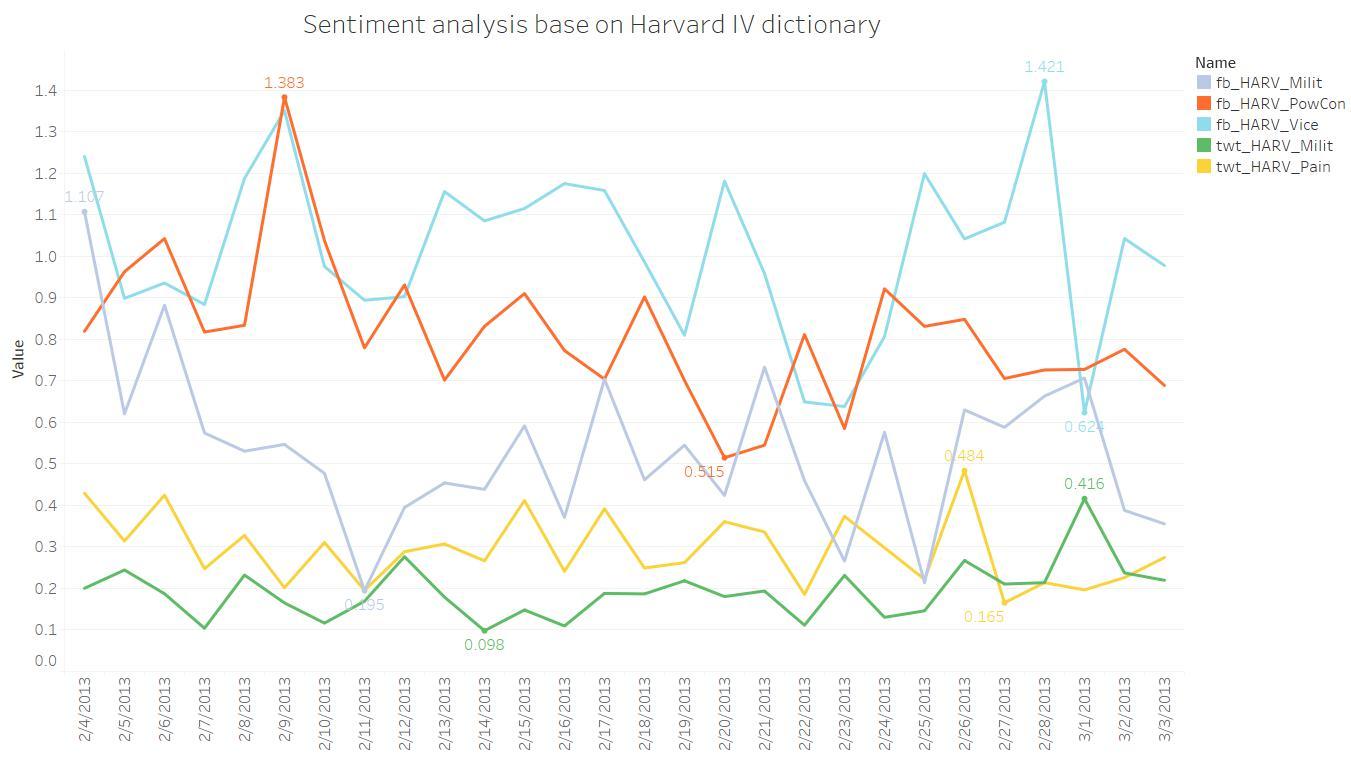


Figure : Line chart of sentiment analysis based on Harvard IV dictionary

Figure 1 is a line chart demonstrates sentiment results in five dimensions base on Harvard IV dictionary and data from Facebook and Twitter used to predict number of conflicts. As legend shows, each line represents one category of dictionary with max and min value on it. fb is the abbreviation of Facebook, and twt is that of Twitter. HARV\_Milit represents military matters, HARV\_Vice shows moral disapproval, HARV\_PowCon reveals power conflict, and HARV\_Pain is injury or suffering. They are all equipped with the ability to predict violence and from the chart, news on Facebook is the most drastic reflection in terms of disgrace and moral disapproval. Besides, twt\_HARV\_Milit is mild from day to day.

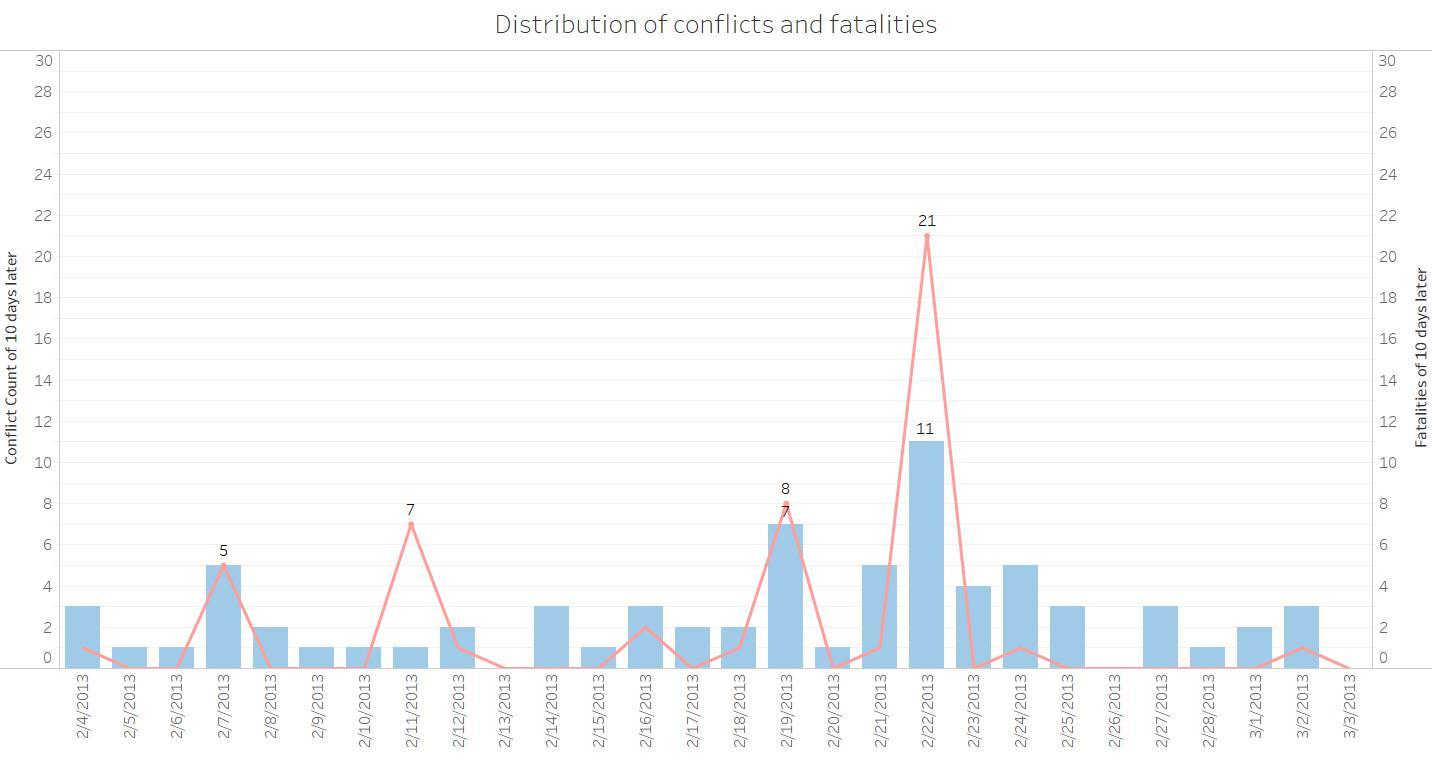


Figure 2 is a combination chart of overall number of conflicts and fatalities for four weeks before the Kenya election, 2013. Blue columns represents the number of conflicts per day with range of 0 to 11, and red line expresses daily fatalities. Conflicts and fatalities both reached peak on 2/22/2013. Riots happened every week, causing many people die in the violence. Condition was the worst in the second week before the election. In order to predict violence, in this project, conflict count of 10 days after election was regarded as dependent variable to predict violence this year.

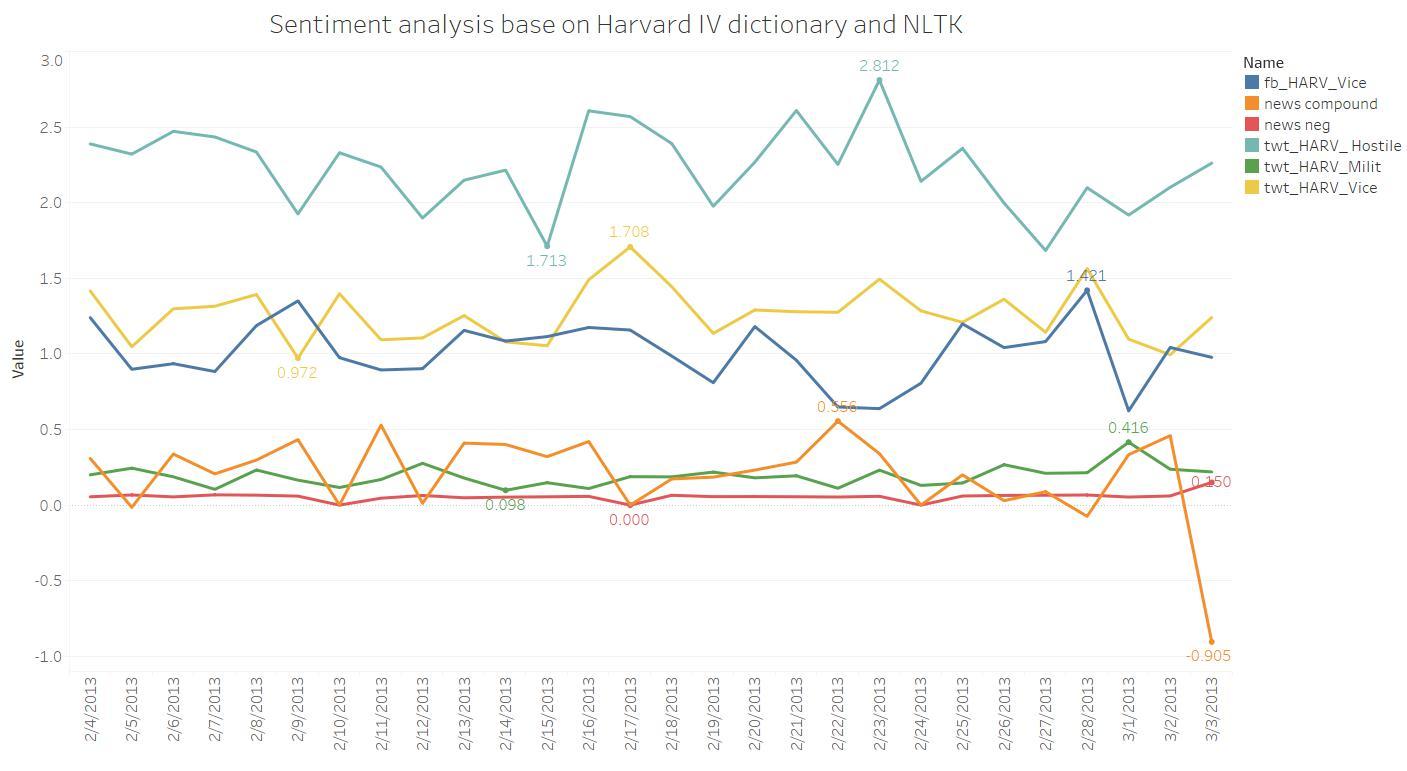


Figure 3 is a line chart representing six types of sentiment based on Twitter, Facebook, news relating to Kenya from Google and Harvard IV dictionary and NLTK package. HARV\_Hostili describes the degree of hostility. News regarding kenya from Google were not very negative since orange and red lines are more stable, but sentiment of news compound dropped to the bottom on the day before the election. Sentiment extracted from Twitter was more radical as three lines described sentiments of Kenyan are over other lines.

**Part B: Violence analysis regarding counties, 2013**

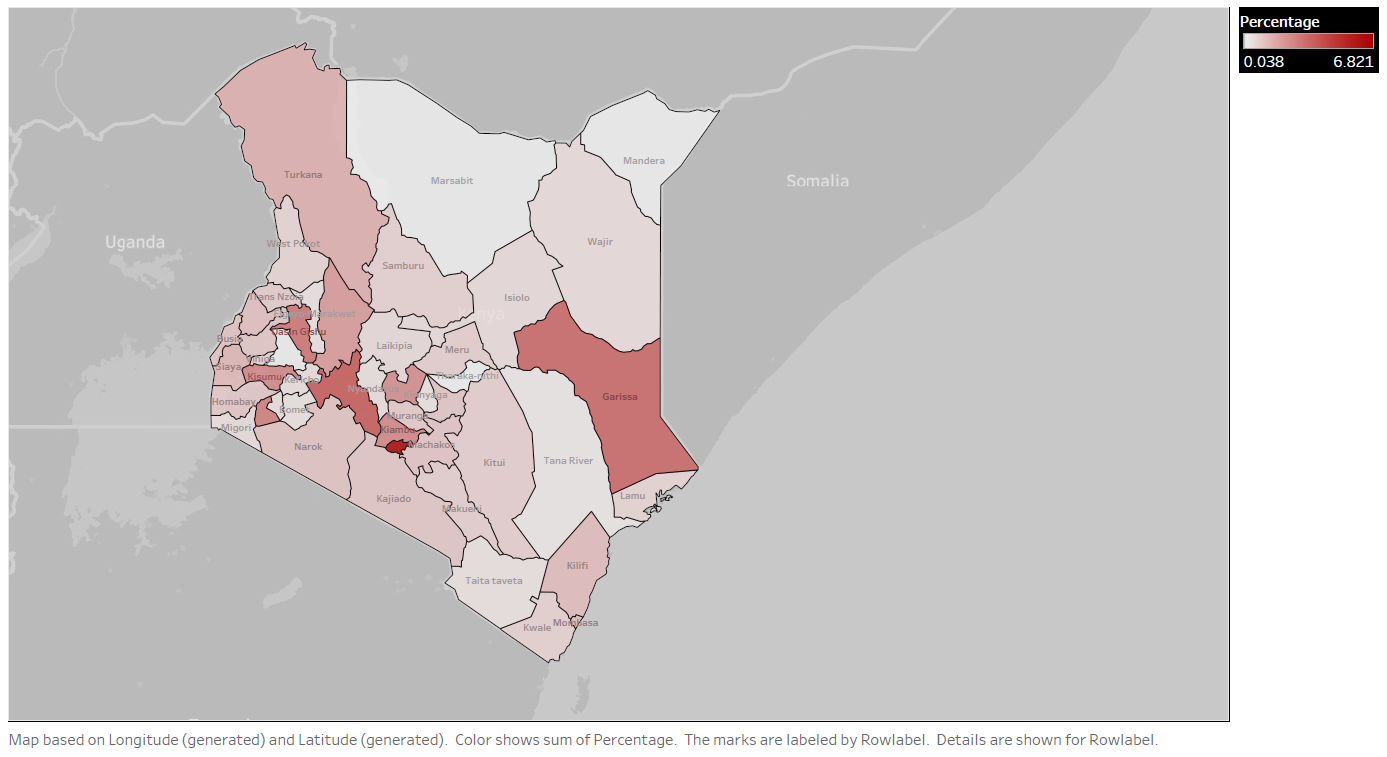
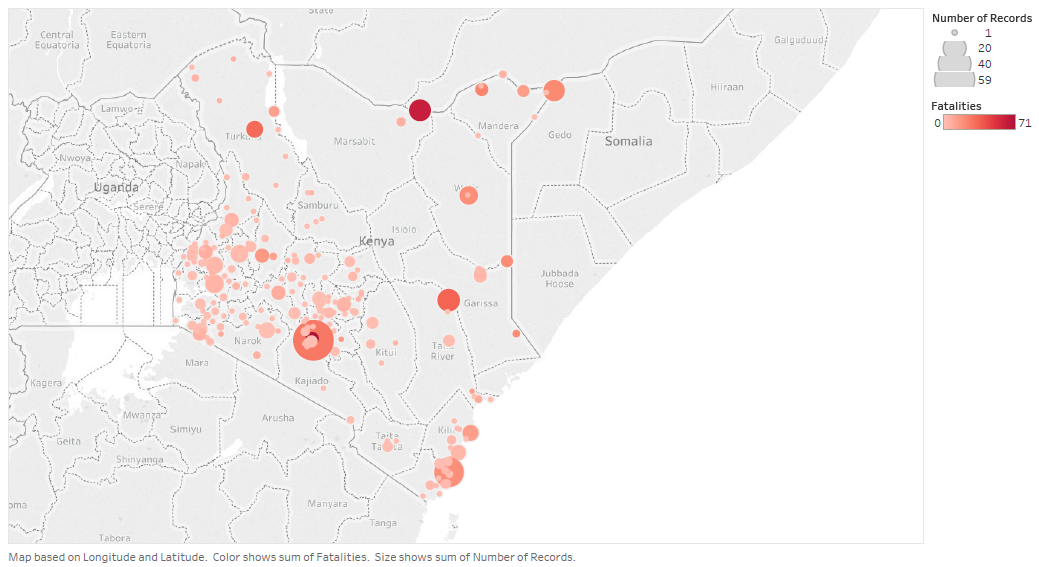


Figure 4

The conflicts in 2013 were distributed in different counties unevenly and were mainly concentrated in Nairobi. Also, Marsabit is the safest place in 2013 since least number of conflicts happened there in 2013.

Figure 5:

Fatalities were mainly in west part of the country. Size of the spot indicated the count of conflicts on that specific location and the color indicated the fatalities of the conflicts. Also, many people died at Nairobi and on the border with Ethiopia. Since the spot at Nairobi is more ‘concentrated’ compared to the others, we can learn that fatalities per conflict is higher than the other places and conflicts were more harsh.

**Part C: Daily violence analysis, four weeks before the election, 2017**

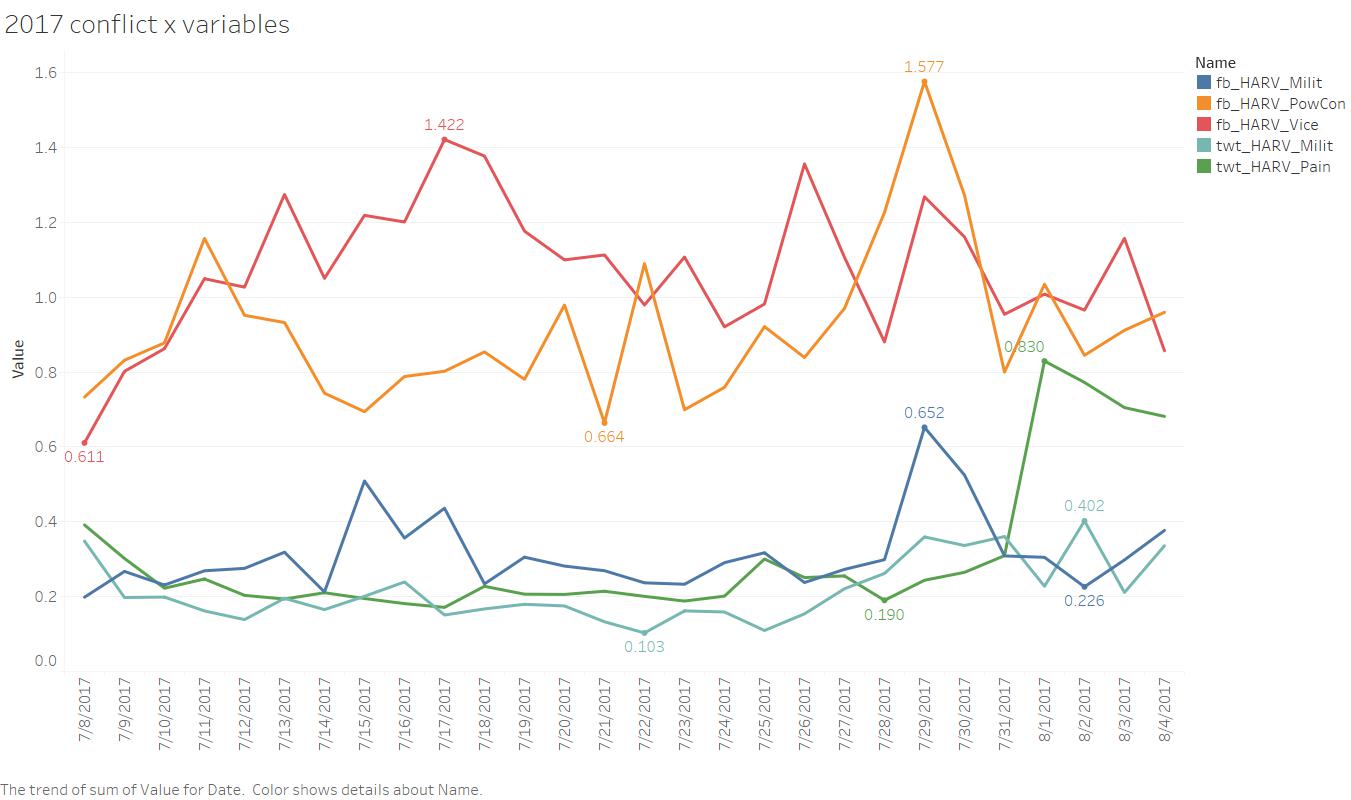


Figure 6 is a line chart demonstrates sentiment results in five dimensions base on Harvard IV dictionary and data from Facebook and Twitter used to predict number of conflicts in 2017. Compared with Figure 1 describing sentiment for 2013, fb\_HARV\_PowCon become the most radical one, fb\_HARV\_Milit is more mild, and both twt\_HARV\_Pain and twt\_HARV\_Milit are still stable.

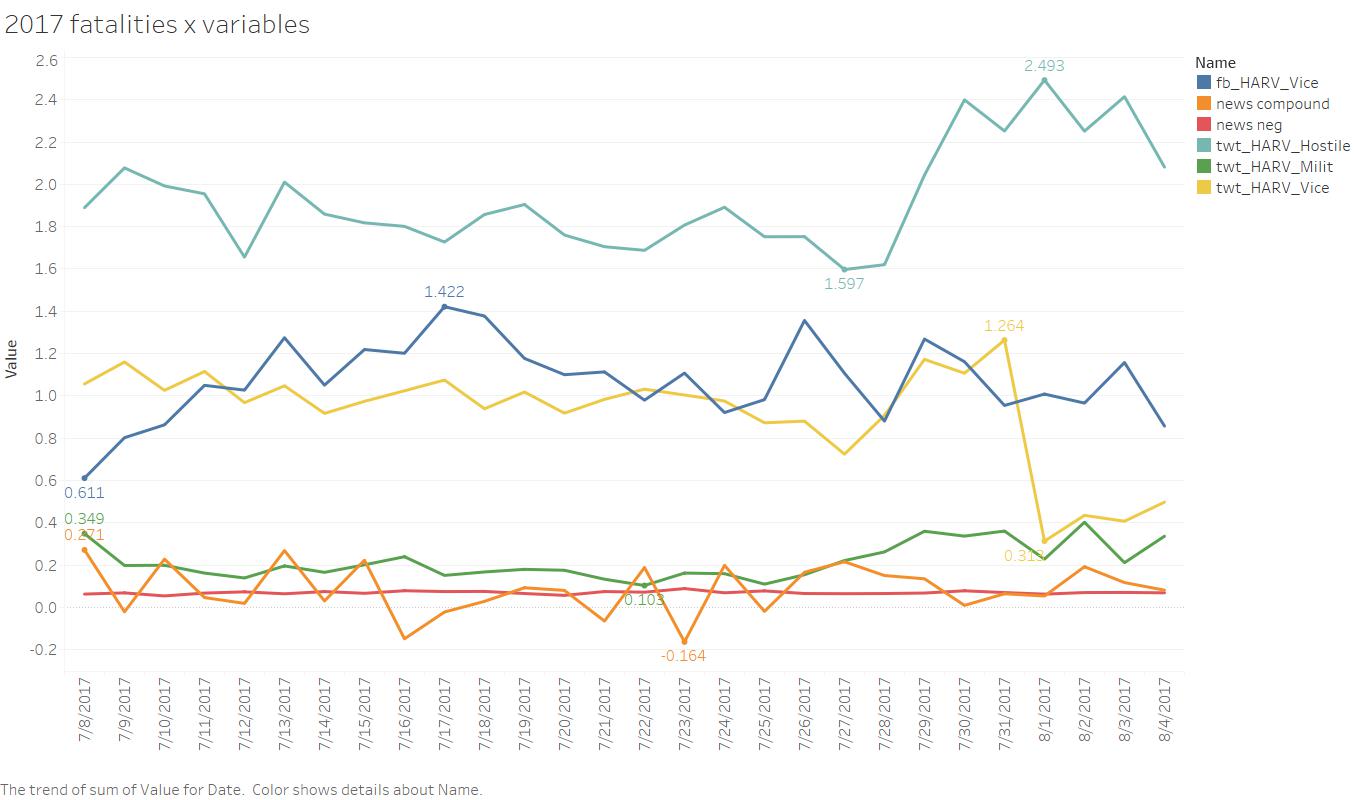


Figure 7 is a line chart representing six types of sentiment based on Twitter, Facebook, news relating to Kenya from Google and Harvard IV dictionary and NLTK package for 2017. Compared with Figure 3, news compound doesn’t drop to the bottom before election, news neg and twt\_HARV\_Milit almost keep the same pattern, and twt\_HARV\_Vice are more smooth and steady.

**Part D: Violence analysis regarding counties, 2017**



Despite owned the highest number of events and root events in the one month before election, Nairobi, the capital district, owned the relative lower percentage of root events. Based on the regression results, Nairobi hence gained relative lower offsetting effect from root indicator, and then lead Nairobi to face the risk of conflicts.