SENTIMENT ANALYSIS AND

TOPIC MODELLING FOR

TWITTER DATA

**1.GAURAV VIVEK KOLEKAR**

**2.VENKATESH**

**3.POOJA LAHOTI**

**INTRODUCTION**

**Background:** Donald J. Trump became the 45th President of the United States on January 20, 2017. Since the inauguration, President Trump is actively initiating new policies and conversations, which generate active conversations in the social media. As Data Scientists, we want to take advantage of this opportunity by using text-mining approaches to conduct social media analytics.

**Hypothesis:**

Our main hypothesis is that the polarity of the tweets will reflect result of the Presidential Election. Majority of tweets from **Blue states** will be **Negative**, majority of the tweets from **Red states** will be **Positive** tweets, Positive tweets would slightly greater than negative tweets in **Swing states**. We have reached such a hypothesis because we are collecting data quite early in the president’s tenure, and it takes some time for people to change their opinion.

In our experiment, we have focused on the following regions:

* Region 1: Florida
* Region 2: California, Oregon, Washington, and Nevada
* Region 3: Texas, Oklahoma
* Region 4: Pennsylvania to Maine
* Region 5: Missouri ,Indiana and Ohio

**Project Overview:** This project is to analyze people’s sentiment and topics about the new administration. We will use Twitter API to collect tweets about Trump. Then we will conduct sentiment analysis to measure how positive or negative the collected tweets are, which can be an indirect measure of Trump’s approval. Next, to see what kinds of topics are discussed related to the new president, we will create word clouds and conduct topic modeling on the collected tweets. To see the geographic variation in opinions, we will collect tweets from 5 different states and conduct the above three analyses to conclude the result.

We have collected 1k tweets from the above mentioned 5 regions as well we conducted the above analyses on 10k tweets collected irrespective location with “TRUMP” as the keyword.

**WORKFLOW**

**Data Collection:**  We have collected 1k tweets using twitter APIs (Tweepy and Twython) from the above mentioned 5 locations using **location (Bounding box coordinates of location) as filter** and **keyword “trump”** in the if condition before storing it to the JSON files. We have also used conditions like the **language of tweet must be English** and **retweeted must be false**.

**Extracting Data from JSON Files:** Once thedata is collected, now we had to extract tweets from the JSON file as the files had many more attributes with it. We have extracted tweets i.e.,

the “text” attribute from JSON file of the raw data and stored in another JSON file which had only tweets (extracted\_Tweets.JSON).

**Sentiment Analysis:**  Now before performing Sentiment Analysis, we cleaned the data using (clean\_tweet) function which removes all the unnecessary symbols (@, #..). We used TextBlob to perform Sentiment analysis and matplotlib to plot the graphs for subjectivity and polarity where X-axis is the score and Y-axis is the tweet count . The polarity score is a float within the range [-1.0, 1.0]. The subjectivity is a float within the range [0.0, 1.0] where 0.0 is very objective and 1.0 is very subjective. We have calculated the average of Subjectivity and Polarity which is indicated by dashed red line in the graphs. We have attached graphs for all the 5 regions as well for 10k tweets which we have collected irrespective of geographical location.

**Word Cloud:**  Before feeding the data in the word cloud, we had to clean data, remove stop words and do stemming. To remove stop words, we used NLTK corpus of stop words and added few more custom stop words (trump, https,http,amp..). We created a special list in which we included the stop words not included in the stop words corpus and removed them. We used porter and Lancaster stemmers to do stemming to the data. Once the above process was done we have given the data into the word cloud module to generate the Word cloud for all the regions.

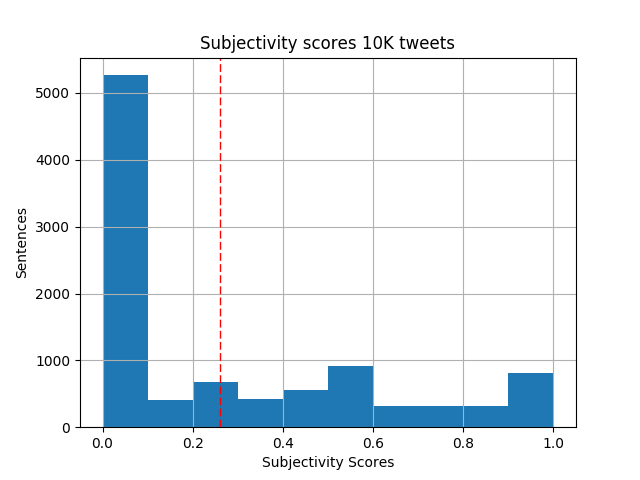
**Topic Modelling:**  It is the technique which gives the approximate topics using the top words.

We have used decomposition and LDA. It (LDA) is a probabilistic model capable of expressing uncertainty about the placement of topics across texts and the assignment of words to topics. Before performing topic modelling also, we cleaned data and removed stop words. We varied the number of topics and number of top words for both the models to get the best topic model for all the regions.

**RESULTS**

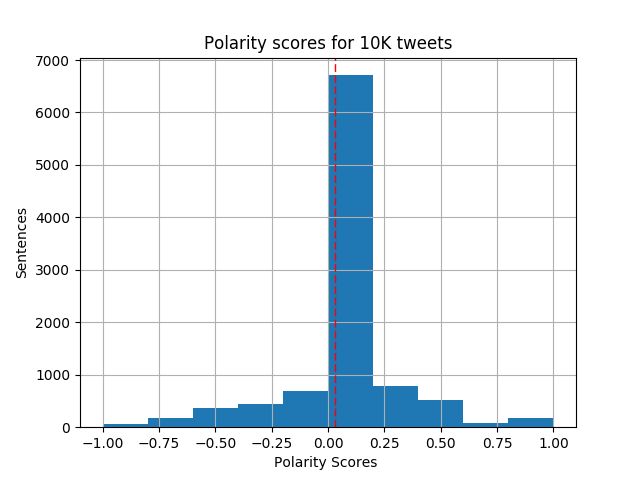
1. **Without any Geographical Location**

**Subjectivity and Polarity plots** for just the keyword without any geographical restrictions



**Figure 1: Subjectivity scores for 10K tweets**

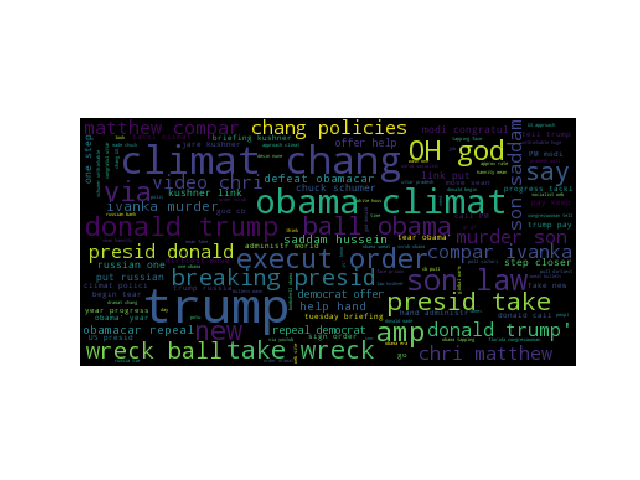
Mean of Subjectivity scores for 10K tweets: **0.260565072067**



**Figure 2: Polarity scores for 10K tweets**

Mean of Polarity scores for 10K tweets: **0.028861495323**

**Word Cloud for 10K tweets**:



**Figure 3: Word Cloud for 10K tweets**

**Topic Modelling Results for 10K tweets:**

The best topic modelling we found were for following parameter:

Number of topics: **7**

Number of top words: **5**

**Result NMF Decomposition:**

Topic 0: climate obama change policies wrecking

Topic 1: compares matthews saddam murderous sons

Topic 2: god oh hannity chuck schumer

Topic 3: kushner puts closer russians briefing

Topic 4: repeal helping hand offer obamacare

Topic 5: florida congresswoman pay tells trump

Topic 6: donald modi calls pm congratulate

**Result of LDA:**

Topic #1 (0, u'0.010\*"climate" + 0.009\*"one" + 0.008\*"Obama" + 0.007\*"Obama\'s" + 0.007\*"years" + 0.007\*"change" + 0.006\*"progress" + 0.006\*"tearing" + 0.006\*"begins" + 0.006\*"tackling"')

Topic #2 (1, u'0.006\*"Russian" + 0.006\*"Trump\'s" + 0.005\*"#Trump" + 0.005\*"lose" + 0.005\*"Want" + 0.005\*"Click" + 0.005\*"weight" + 0.005\*"fast?" + 0.004\*"Nunes" + 0.004\*"set"')

Topic #3 (2, u'0.006\*"Democrats" + 0.006\*"Obamacare" + 0.006\*"helping" + 0.006\*"defeat" + 0.006\*"hand" + 0.006\*"repeal," + 0.006\*"offer" + 0.005\*"administration" + 0.005\*"World" + 0.004\*"Socialist"')

Topic #4 (3, u'0.008\*"Chris" + 0.008\*"sons" + 0.008\*"Matthews" + 0.008\*"compares" + 0.008\*"murderous" + 0.008\*"VIDEO:" + 0.007\*"SADDAM" + 0.007\*"HUSSEIN!" + 0.007\*"Ivanka" + 0.006\*"undo"')

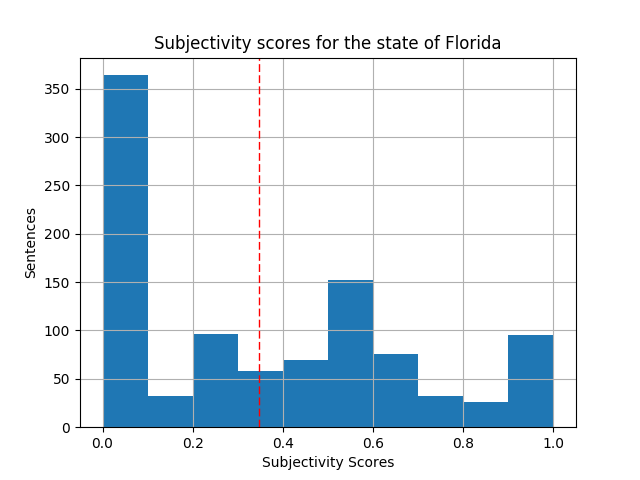
Topic #5 (4, u'0.011\*"OH" + 0.011\*"GOD!" + 0.009\*"Trump:" + 0.009\*"Keep" + 0.009\*"Florida" + 0.008\*"Pay" + 0.008\*"Congresswoman" + 0.008\*"Tells" + 0.007\*"Chuck" + 0.007\*"HUGE!"')

Topic #6 (5, u'0.008\*"Made" + 0.003\*"Clintons" + 0.002\*"rails" + 0.002\*"NATO" + 0.002\*"via" + 0.002\*"Donald" + 0.002\*"#Trump" + 0.002\*"&amp;" + 0.002\*"Power" + 0.002\*"\'hoax,\'"')

Topic #7 (6, u'0.010\*"take" + 0.010\*"BREAKING:" + 0.010\*"Change" + 0.009\*"WRECKING" + 0.009\*"BALL" + 0.009\*"Obama\u2019s" + 0.009\*"policies!\u2026" + 0.009\*"Climate" + 0.008\*"President" + 0.007\*"change"')

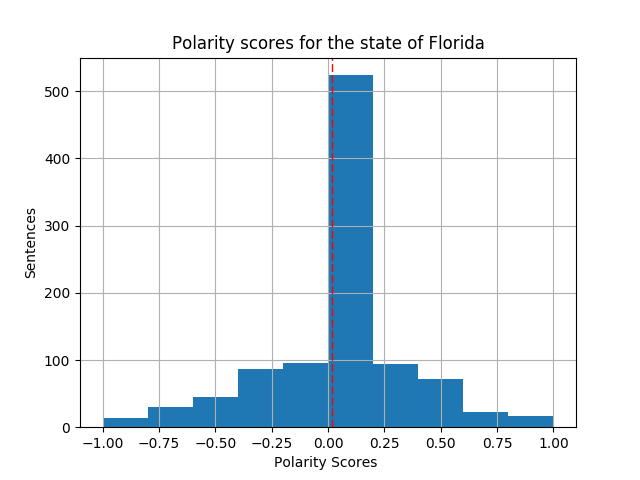
**REGION 1: Florida**

The following are the **Subjectivity and Polarity plots for Region 1 (Florida)**:



**Figure 4: Subjectivity score for the state of Florida**

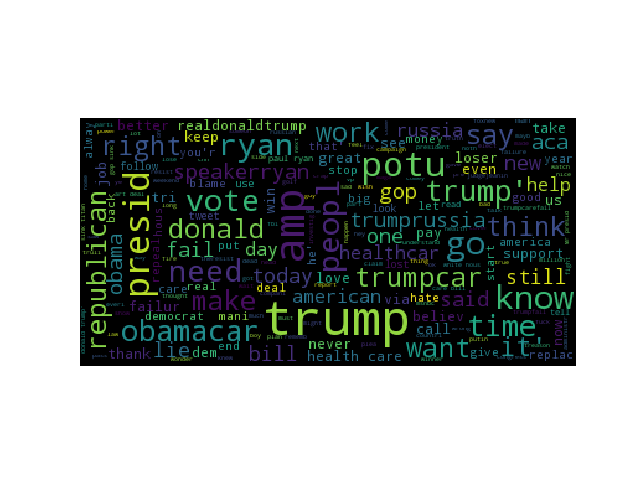
Mean of Subjectivity score for Region 1: **0.347396305223**



**Figure 5: Polarity scores for the state of Florida**

Mean of Polarity score for Region 1: **0.0198090186203**

**Word Cloud for Region 1 (Florida):**



**Figure 6: Word Cloud for Region 1 (Florida)**

**Topic Modelling Results for Region 1 (Florida)**:

The best topic modelling we found were with the following parameters:

Number of topics: **7**

Number of top words: **5**

**Result of NMF decomposition:**

Topic 0: trump failure obamacare people care

Topic 1: trumpcare dead care obamacare health

Topic 2: gop like latest titanic article

Topic 3: president ryan mr paul need

Topic 4: article said right repeal russia

Topic 5: donald presidency joke know hate

Topic 6: potus speakerryan help time people

**Result of LDA:**

Topic #1 (0, u'0.002\*"Sinking" + 0.002\*"GOP..." + 0.002\*"article...." + 0.002\*"Titanic" + 0.002\*"like" + 0.002\*"latest" + 0.002\*"Donald" + 0.002\*"Obamacare" + 0.002\*"Trumpcare" + 0.002\*"TheResistance"')

Topic #2 (1, u'0.002\*"all;" + 0.002\*"companies" + 0.001\*"can\'t" + 0.001\*"still" + 0.001\*"SpeakerRyan" + 0.001\*"pay" + 0.001\*"people" + 0.001\*"help" + 0.001\*"many" + 0.001\*"Trumpcare"')

Topic #3 (2, u'0.002\*"Donald" + 0.002\*"Trump\'s" + 0.002\*"people" + 0.002\*"love" + 0.002\*"healthcare" + 0.002\*"Hospitality" + 0.002\*"Miami," + 0.002\*"President" + 0.002\*"Trump" + 0.002\*"&amp;"')

Topic #4 (3, u'0.003\*"Ryan" + 0.002\*"President" + 0.002\*"Donald" + 0.002\*"president" + 0.001\*"via" + 0.001\*"need" + 0.001\*"know" + 0.001\*"Service" + 0.001\*"like" + 0.001\*"take"')

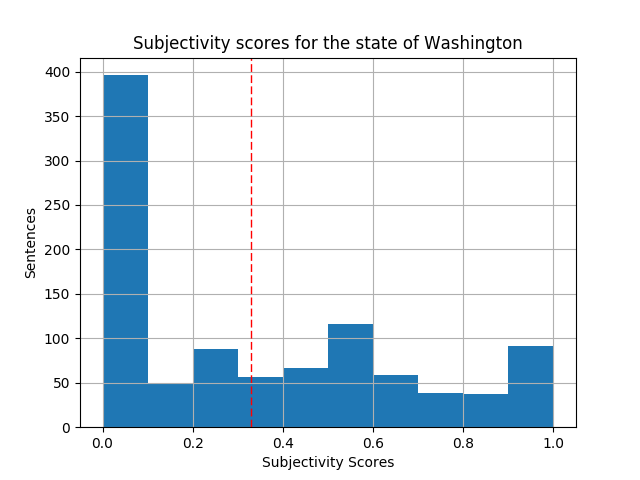
Topic #5 (4, u'0.003\*"pay" + 0.002\*"always" + 0.002\*"people" + 0.002\*"use" + 0.002\*"President" + 0.002\*"POTUS" + 0.002\*"hate" + 0.002\*"want" + 0.002\*"anymore" + 0.001\*"still"')

Topic #6 (5, u'0.002\*"POTUS" + 0.002\*"ample;" + 0.002\*"Thanks" + 0.002\*"think" + 0.002\*"Presidency" + 0.002\*"President" + 0.002\*"Joke" + 0.001\*"Already" + 0.001\*"know" + 0.001\*"House"')

Topic #7 (6, u'0.002\*"start" + 0.002\*"great" + 0.002\*"Trump\'s" + 0.002\*"investigation" + 0.001\*"vote" + 0.001\*"Obama" + 0.001\*"say" + 0.001\*"ass;" + 0.001\*"WATCH:" + 0.001\*"Corn"')

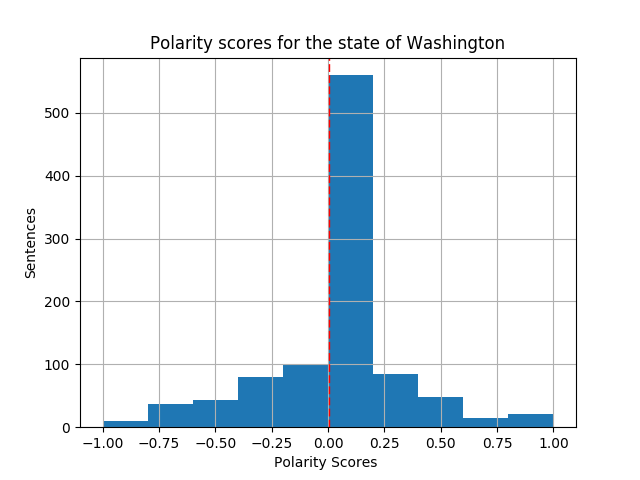
**REGION 2: California, Oregon, Washington, Nevada**

The following are the **Subjectivity and Polarity plots for Region 2:**



**Figure 7: Subjectivity scores for Region 2**

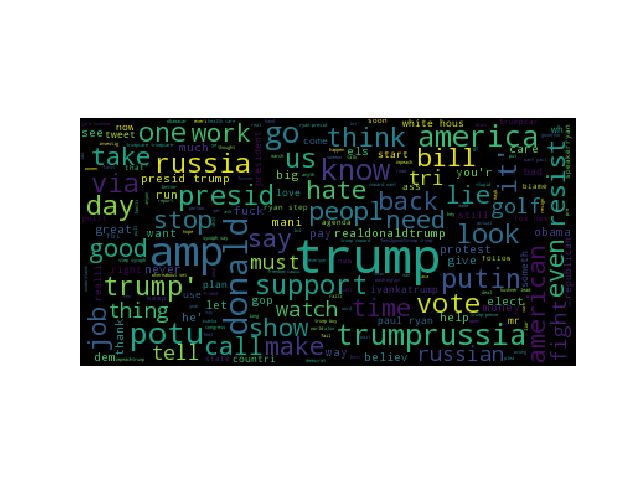
Mean of Subjectivity scores for Region 2: **0.329016246098**



**Fig 8: polarity scores for region 2**

Mean of Polarity scores for Region 2: **0.0057037636627**

**Word Cloud for Region 2**:



**Fig 9: Word Cloud for Region 2**

**Topic Modelling Results for Region 2:**

The best topic modelling we found were with the following parameters:

Number of topics: **7**

Number of top words: **5**

Result of **NMF Decomposition**:

Topic 0: trump russia know putin presidency

Topic 1: ryan like don paul say

Topic 2: donald dick suck brendonurie fuck

Topic 3: amp trumprussia putin resist russiagate

Topic 4: news fox alert critics whereabouts

Topic 5: kushner lead white house government

Topic 6: president potus golf america merkel

**Result of LDA:**

Topic #1 (0, u'0.003\*"Trump\'s" + 0.002\*"alert" + 0.002\*"whereabouts" + 0.002\*"critics" + 0.002\*"Fox" + 0.002\*"set" + 0.002\*"off\u201d" + 0.002\*"News" + 0.002\*"\u26a1\ufe0f" + 0.001\*"\u201cA"')

Topic #2 (1, u'0.001\*"get" + 0.001\*"&amp;" + 0.001\*"#Trump" + 0.001\*"PRESIDENT" + 0.001\*"outrage?" + 0.001\*"believe" + 0.001\*"@POTUS" + 0.001\*"love" + 0.001\*"wrong" + 0.001\*"agenda"')

Topic #3 (2, u'0.002\*"\u270a\u270a\u270a" + 0.002\*"Resist" + 0.001\*"Trump\'s" + 0.001\*"#TrumpRussia" + 0.001\*"Trump." + 0.001\*"going" + 0.001\*"Donald" + 0.001\*"fight" + 0.001\*"#IMWithHer" + 0.001\*"get"')

Topic #4 (3, u'0.002\*"Merkel" + 0.002\*"ever" + 0.002\*"gave" + 0.001\*"\'invoice\'" + 0.001\*"alleged" + 0.001\*"NATO" + 0.001\*"Former" + 0.001\*"Contradict" + 0.001\*"Claims" + 0.001\*"Miss"')

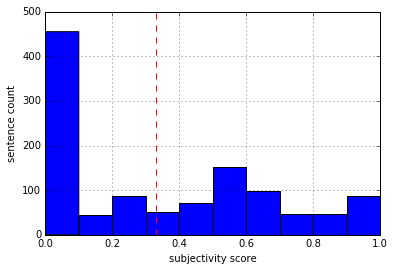
Topic #5 (4, u'0.003\*"Paul" + 0.003\*"Ryan" + 0.003\*"Don\'t" + 0.003\*"Trump," + 0.003\*"like" + 0.003\*"President" + 0.002\*"Coward:" + 0.002\*"Ryan," + 0.002\*"Sycophant" + 0.002\*"Step"')

Topic #6 (5, u'0.002\*"like" + 0.002\*"&amp;" + 0.001\*"#TrumpRussia" + 0.001\*"would" + 0.001\*"#Trumpleaks" + 0.001\*"SOMETHING" + 0.001\*"MAD" + 0.001\*"#WomensMarch\u2026" + 0.001\*"HELL" + 0.001\*"say"')

Topic #7 (6, u'0.002\*"@POTUS" + 0.002\*"&amp;" + 0.002\*"Jared" + 0.002\*"Donald" + 0.001\*"Kushner" + 0.001\*"House" + 0.001\*"via" + 0.001\*"force" + 0.001\*"task" + 0.001\*"common,"')

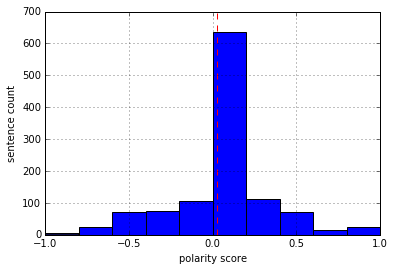
**REGION 3: TEXAS and OKLAHOMA**

The following are the **Subjectivity and Polarity plots for Region 3:**



**Figure 10: Subjectivity of Texas and Oklahoma Region**

**Subjectivity Mean for Region 3: 0.33095165801**



**Figure 11: Polarity of Texas and Oklahoma Region**

**Polarity Mean for Region 3: 0.0314546918844**

**Word Cloud for Region 3:**



**Figure 12: Word Cloud without Stemming**



**Figure 13: Word Cloud with Stemming**

**TOPIC MODELLING for Region 3:**

**Number of Topics : 7**

**Number of top words : 5**

**Results from NMF Decomposition:**

Topic 0: good hes healthcare president fuck

Topic 1: obamacare ryan aca cute reform

Topic 2: day jazz demand cover trying

Topic 3: pausethispresidency act 2017s beaner destroying

Topic 4: great maga buildthatwall abbott wrote

Topic 5: needed investigation reminder president selectcommittee

Topic 6: wiretap surveillance ask bs monitoring

**Results of LDA model :**

Topic #1 (0, u'0.001\*"true" + 0.001\*"DNC" + 0.001\*"verify" + 0.001\*"CrowdStrike" + 0.001\*"claimed" + 0.001\*"attitude" + 0.001\*"General" + 0.001\*"plus" + 0.001\*"Flynn" + 0.001\*"destroy"')

Topic #2 (1, u'0.001\*"happened" + 0.001\*"bill" + 0.001\*"picking" + 0.001\*"Col" + 0.001\*"Peters" + 0.001\*"day" + 0.001\*"really" + 0.001\*"didnt" + 0.001\*"start" + 0.001\*"What"')

Topic #3 (2, u'0.001\*"hurting" + 0.001\*"Benefits" + 0.001\*"Always" + 0.001\*"1st" + 0.001\*"cutting" + 0.001\*"focus" + 0.001\*"immigration" + 0.001\*"kids" + 0.001\*"voters" + 0.001\*"needs"')

Topic #4 (3, u'0.001\*"cuddle" + 0.001\*"splendor" + 0.001\*"wuddle" + 0.001\*"little" + 0.001\*"needed" + 0.001\*"SelectCommittee" + 0.001\*"absolutely" + 0.001\*"intended" + 0.001\*"haircut" + 0.001\*"Damn"')

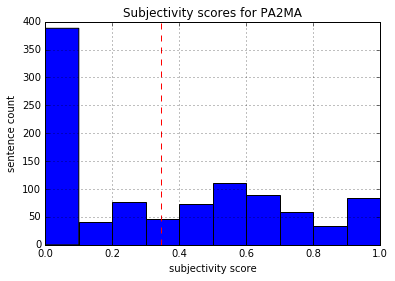
Topic #5 (4, u'0.001\*"briefing" + 0.001\*"approach" + 0.001\*"Nunez" + 0.001\*"asked" + 0.001\*"recuse" + 0.001\*"Who" + 0.001\*"county" + 0.001\*"reminder" + 0.001\*"Immigrants" + 0.001\*"Hillary"')

Topic #6 (5, u'0.001\*"charity" + 0.001\*"fuk" + 0.001\*"flexing" + 0.001\*"con" + 0.001\*"loser" + 0.001\*"jerksecond" + 0.001\*"BOBBcolluded" + 0.001\*"stupidDems" + 0.001\*"kkk" + 0.001\*"familyhis"')

Topic #7 (6, u'0.001\*"connected" + 0.001\*"ask" + 0.001\*"surveillance" + 0.001\*"BS" + 0.001\*"IT" + 0.001\*"blames" + 0.001\*"Tomi" + 0.001\*"outstanding" + 0.001\*"poormiddle" + 0.001\*"3PM"')

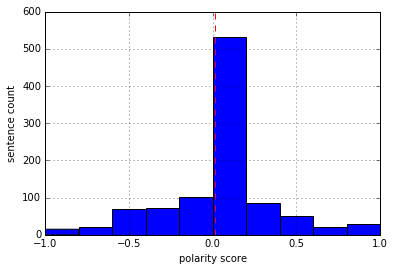
**Region 4: Pennsylvania to Maine**

The following are the **Subjectivity and Polarity plots for Region 4:**



**Figure 14: Subjectivity of Pennsylvania to Maine**

**Subjectivity Mean for Region 4: 0.344399151616**



**Figure 15: Polarity of Pennsylvania to Maine**

**Polarity Mean for Region 4: 0.0114431215427**

**Word Cloud for Region 4:**



**Figure 16: Word Cloud for Region 4 with Stemming**

**Topic Modelling for Region 4:**

**Number of Topics : 7**

**Number of top words : 5**

**Results of NMF decomposition:**

Topic 0: people care good dont ryan

Topic 1: people resist make russia new

Topic 2: asset liabilityfakepresident entire president nsa

Topic 3: people dont care need want

Topic 4: dont new toms thought god

Topic 5: asset entire liabilityfakepresident obamacare president

Topic 6: obama president asset think nsa

**Results of LDA model:**

Topic #1 (0, u'0.001\*"till" + 0.001\*"RESISTANCE" + 0.001\*"Report" + 0.000\*"calls" + 0.000\*"joke" + 0.000\*"cost" + 0.000\*"communist" + 0.000\*"NoAHCA" + 0.000\*"quote" + 0.000\*"FakePOTUS"')

Topic #2 (1, u'0.001\*"loser" + 0.001\*"KillTheBill" + 0.001\*"awful" + 0.000\*"tough" + 0.000\*"refer" + 0.000\*"discover" + 0.000\*"room" + 0.000\*"asks" + 0.000\*"ready" + 0.000\*"paperback"')

Topic #3 (2, u'0.001\*"EPA" + 0.001\*"directionChina" + 0.001\*"NSA" + 0.001\*"supreme" + 0.001\*"lead" + 0.001\*"whistleblower" + 0.001\*"Binney" + 0.001\*"states" + 0.001\*"unit" + 0.001\*"wiretapped"')

Topic #4 (3, u'0.001\*"Toms" + 0.001\*"pussy" + 0.001\*"tcot" + 0.001\*"ive" + 0.001\*"subtweets" + 0.001\*"bless" + 0.001\*"tool" + 0.001\*"bit" + 0.000\*"Jersey" + 0.000\*"River"')

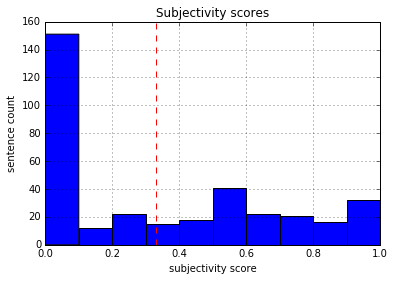
Topic #5 (4, u'0.002\*"asset" + 0.001\*"liabilityfakepresident" + 0.001\*"entire" + 0.001\*"liabilityFake" + 0.001\*"confused" + 0.001\*"NSA" + 0.001\*"foxnews" + 0.001\*"globalist" + 0.001\*"lavar" + 0.001\*"cat"')

Topic #6 (5, u'0.001\*"total" + 0.001\*"spirited" + 0.001\*"city" + 0.001\*"rally" + 0.001\*"bc" + 0.000\*"home" + 0.000\*"SAD" + 0.000\*"center" + 0.000\*"accidentally" + 0.000\*"Eric"')

Topic #7 (6, u'0.001\*"bet" + 0.000\*"wont" + 0.000\*"passes" + 0.000\*"win" + 0.000\*"truck" + 0.000\*"whos" + 0.000\*"decide" + 0.000\*"verbal" + 0.000\*"feckless" + 0.000\*"fear"')

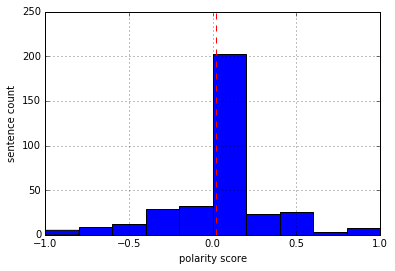
**Region 5: Missouri, OHIO, Indiana**

The following are the subjectivity and polarity for region 5:



**Figure 17: Subjectivity for Missouri to Indiana**

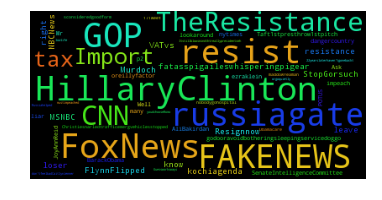
**Subjectivity Mean for Region 5: 0.331871087088**



**Figure 18: Polarity for Missouri to Indiana**

**Polarity Mean for Region 5: 0.0188225497981**

**Word Cloud for Region 5:**



**Figure 18: Word Cloud for region 5**

**Topic Modelling for Region 5:**

Number of Topics: 7

Number of Top words: 3

**Results for NMF Decomposition:**

Topic 0: ass country need

Topic 1: tax pig ass

Topic 2: really russia ass

Topic 3: hillaryclinton people doggo

Topic 4: inelearn loser Flynn

**Results of LDA Model:**

Topic #1 (0, u'0.001\*"pig" + 0.001\*"changing" + 0.001\*"tax" + 0.001\*"Taft" + 0.001\*"p2" + 0.001\*"kochi" + 0.001\*"vs" + 0.001\*"though" + 0.001\*"Dems2018" + 0.001\*"indeed"')

Topic #2 (1, u'0.001\*"could" + 0.001\*"hearing" + 0.001\*"prez" + 0.001\*"USA" + 0.001\*"obvious" + 0.001\*"con" + 0.001\*"Nunes" + 0.001\*"shape" + 0.001\*"tcot" + 0.001\*"traitors"')

Topic #3 (2, u'0.001\*"people" + 0.001\*"supporter" + 0.001\*"look" + 0.001\*"thing" + 0.001\*"way" + 0.001\*"coming" + 0.001\*"also" + 0.001\*"liar" + 0.001\*"talking" + 0.001\*"need"')

Topic #4 (3, u'0.001\*"people" + 0.001\*"supporter" + 0.001\*"look" + 0.001\*"thing" + 0.001\*"way" + 0.001\*"also" + 0.001\*"coming" + 0.001\*"liar" + 0.001\*"talking" + 0.001\*"need"')

Topic #5 (4, u'0.002\*"INeLearn" + 0.001\*"HillaryClinton" + 0.001\*"go" + 0.001\*"NEWS" + 0.001\*"FAKE" + 0.001\*"LOSER" + 0.001\*"resignnow" + 0.001\*"potato" + 0.001\*"omg" + 0.001\*"Flip"')

**INSIGHTS**

1. Number of positive and Negative tweets for each region (per 1000)

where positive tweet are all tweets polarity greater than 0 and negative tweets are less than 0

**For region 1:**

Positive tweets: 323

Negative tweets: 270

**For region 2:**

Positive tweets: 292

Negative tweets: 272

**For region 3:**

Positive tweets: 352

Negative tweets: 279

**For region 4:**

Positive tweets: 302

Negative tweets: 279

**For region 5:**

Positive tweets: 109

Negative tweets: 87

2. Number of neutral tweets (where we considered neutral tweet is between -0.5 and 0.5)

Region 1: 827

Region 2: 855

Region 3: 971

Region 4: 831

Region 5: 295

3. Average followers of people who tweet positively (i.e polarity between 0.5 and 1) about the keyword

Region 1: 2920.4

Region 2: 1028.88636364

Region 3: 997.6

Region 4: 1019.20754717

Region 5: 1668.0

4. Average followers of people who tweet negatively (i.e polarity between -1 and -0.5) about the keyword

Region 1: 1122.0

Region 2: 997.555555556

Region 3: 886.5

Region 4: 1164.09090909

Region 5: 986.466666667

5. **Insights from Topic modelling:**

**a. Common topics of discussion in all 5 regions:**

**1.Obama Care**

**2.Healthcare**

**3.POTUS (President of United States)**

**4. Russia,#TrumpRussia , #Russiagate ,Putin**

**5.Speaker Ryan , Mr.Paul**

**6.Nunes**

**b. Unique Topics among the regions:**

**1. Region 2 and Region 5:** Though there is a geographical distance between 2 regions **Foxnews** is a common topic

**2. Region 2: Merkel** visit to WH was unique topic of interest **.**

**3.Region 5 and Region 3:** There is no much geographical distance between 2 regions, s**o #MAGA (Make America Great Again) ,HillaryClinton** are common topics.

**4.Region 4: Whistleblower Bill binney, #NoAHCA (No America Health Care Act)**

**5. Region 3: Immigrants , Buildthewall , Flynn** are unique topics as we suppose Mexico is near **.**

**6. Region 5: #Dems2018 Traitors** is most discussed.

**7. Region 1** do not any unique topics**.**

**V. CONCLUSION**

We didn’t get to see as many variation as we were expecting. Maybe because data is extremely less for us to prove or disprove our main hypothesis

**VI. REFERENCES**

1. <https://stackoverflow.com/>

2. <http://www.geeksforgeeks.org/twitter-sentiment-analysis-using-python/>

3.Class Code