PROJECT REPORT-SOCIAL MEDIA ANALYTICS

Team Members:

Akanksha Shrivastava

Gulrej Hamirani

Vijetha Shenoy

PROJECT REPORT-SOCIAL MEDIA ANALYTICS

Objective:

The purpose of this project is to conduct Social Media Analytics by using Text Mining approaches to analyse the collected tweets and Trump's influence in selected 5 states and overall sentiment. We have chosen a mix of blue and red states to avoid any bias while deriving our insights.

As per 2016 elections, below are the list of selected states demonstrating the color of states.

California – Blue state (Democratic win over 5%)

Colorado – Blue state (Democratic win under 5%)

Florida – Red state (Republic win under 5%)

New York – Blue state (Democratic win over 5%)

Texas – Red state (Republic win over 5%)

High level Code:
For 10K tweets with trump keyword

- We started by downloading 10K tweets with trump keywords. We did it in chunks due to connection break issues when we run 10k at once. Later, we merged all the chunked tweets into one.
- We did sentiment analysis on tweets. We defined a function to remove all non-ascii characters as well as digits, punctuations using regular expression and plotted the histogram for subjectivity and polarity scores.
- For word cloud, we split the tweets into words and saved them into a list. Ran a function that removed all the stopwords from the tweets. We removed irrelevant words by appending them to the stopwords list.
- After removing stopwords, we performed stemming using SnowballStemmer and plotted word cloud.
- For topic modelling, we have used NMF and LDA. For NMF, the corpus had to be a list of tweets for which we removed all punctuation, stopwords. After doing stemming, we converted it again to list of strings and passed it to corpus and defined number of topics as 10 and printed the topics.
- For LDA the corpus had to be a list of lists so we converted the tweets to a list of words, removed stopwords and again converted to a list of lists and passed it to a corpus, assigned number of topics as 10 and printed the topics.
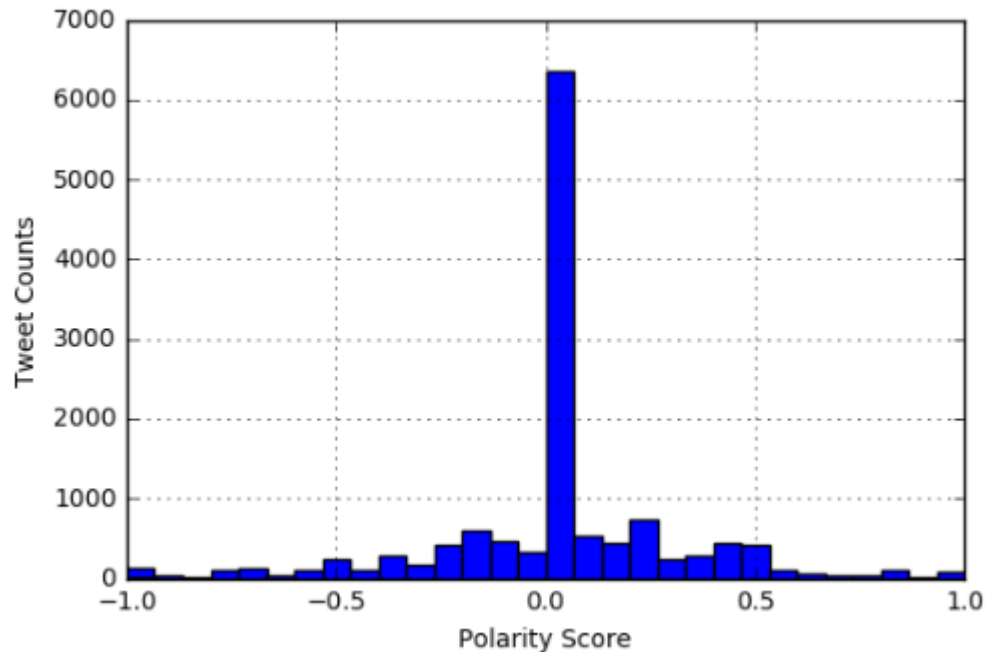
For Location based tweets with trump keywords:
- All the above-mentioned steps were performed except while collecting tweets the location coordinates were passed as an argument i.e. long1, lat1, long2, lat2 which we looked up from Google maps by creating virtual rectangle around the state.
- Also, filtered the tweets based on the "trump" keywords and repeated the same process which we did for the 10K tweets.
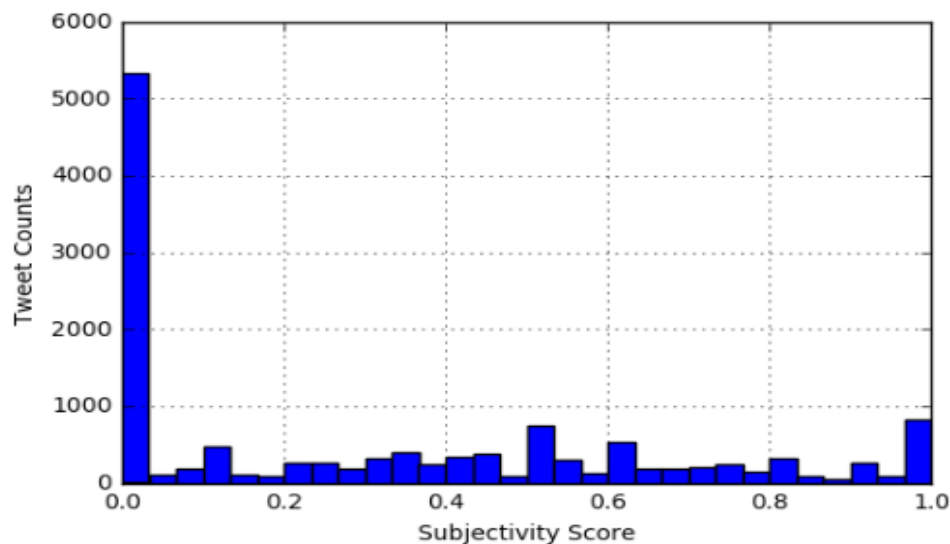
PROJECT REPORT-SOCIAL MEDIA ANALYTICS

Sentiment Analysis of 10k tweets:

Below are polarity and subjectivity scores plotted using histogram along with average Polarity and Subjectivity scores.

Considering the average polarity score, we can say that the average sentiment is slightly above zero indicating sentiment being slightly positive than neutral.


```
Average of Polarity Scores: 0.0160633037994
Average of Subjectivity Scores: 0.311068197702
```

Considering the average subjectivity score, we can say that the average sentiment is 0.3 indicating tweets to be fairly subjective. However, since most of the tweets collected contain sarcasm, slang and possibly spelling mistakes, it is unable to catch the positive or negative sentiment accurately.

PROJECT REPORT-SOCIAL MEDIA ANALYTICS

Word cloud of 10k tweets:

Topic Modeling of 10k tweets:

NMF

```
Topic 0: time mani melania showed ladi donald russian investig presid fbi
Topic 1: ladi showed mani melania time lie twitter russian bank money
Topic 2: lie twitter capabl fbi day tower donald impo purpo support
Topic 3: ladi showed mani melania time putin gop come think advi
Topic 4: donald presid russia lie know fbi news ivanka investig new
Topic 5: day investig think everi sit anoth wh democraci theresistance destroy
Topic 6: suppo russia fbi promi presid agr ivanka mean investig donald
Topic 7: campaign russia suppo crimin russian money propo report donald colleg
Topic 8: russia donald ivanka obama presid suppos fbi putin ralli promis
Topic 9: suppos russia agre promis investigation advis wiretap ivanka america court
Topic 10: th believ hackedtvshow learn west wing fuck donald etribun lawmak
Topic 11: russia donald wiretap children lie ivanka investig obama promis congress
Topic 12: purpo impo th caus peopl fund yemeni uh help aid
Topic 13: wiretap presid ivanka russia congress sb lie news america usa
Topic 14: donald promi advi say agr suppo presid peopl russian vote
Topic 15: capabl tower colleg nomin terror win act breitbart award land
Topic 16: know russia peopl donald lie fbi claim want time opinion
Topic 17: donald obama russia presid ooh white grab tweet lie russian
Topic 18: presid right donald peopl russia investig know say russian admin
Topic 19: donald russian russia man manafort run judg alleg compani laun
```

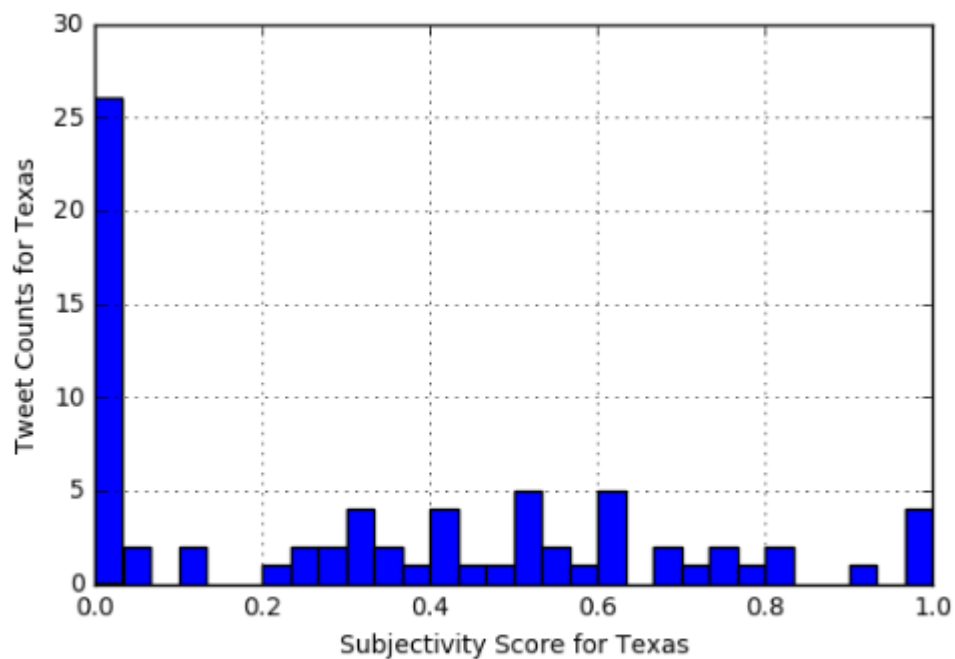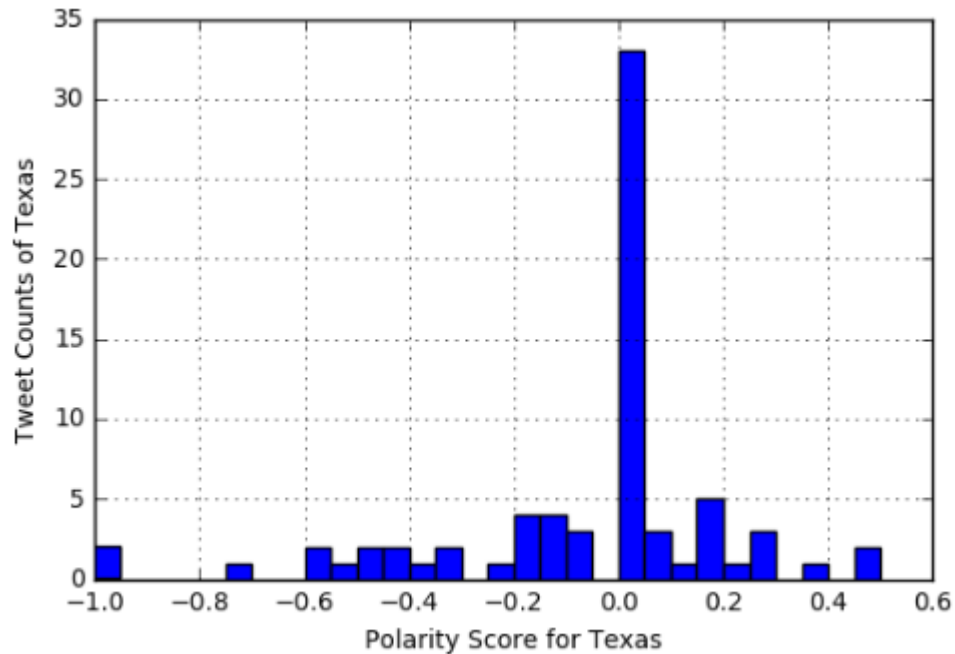# PROJECT REPORT-SOCIAL MEDIA ANALYTICS

LDA

```
[(0,
  u'0.016*"budget" + 0.013*"donald" + 0.010*"thing" + 0.010*"tower" + 0.010*"wiretap" + 0.009*"obama" + 0.008*"senat"
+ 0.007*"never" + 0.006*"indic" + 0.006*"wife"'),
 (1,
  u'0.018*"donald" + 0.014*"budget" + 0.013*"senat" + 0.010*"tower" + 0.009*"wiretap" + 0.007*"never" + 0.007*"thing"
+ 0.006*"obama" + 0.006*"intel" + 0.006*"knew"'),
 (2,
  u'0.013*"donald" + 0.012*"budget" + 0.012*"tower" + 0.012*"obama" + 0.011*"senat" + 0.010*"wiretap" + 0.009*"indic"
+ 0.007*"us" + 0.006*"intel" + 0.006*"lie"'),
 (3,
  u'0.023*"donald" + 0.017*"thing" + 0.017*"knew" + 0.017*"never" + 0.017*"budget" + 0.016*"wife" + 0.011*"senat" +
0.009*"wiretap" + 0.008*"tower" + 0.006*"ban"'),
 (4,
  u'0.014*"budget" + 0.014*"donald" + 0.013*"senat" + 0.012*"wiretap" + 0.010*"tower" + 0.009*"obama" + 0.007*"thing"
+ 0.007*"claim" + 0.006*"never" + 0.006*"intel"'),
 (5,
  u'0.014*"donald" + 0.013*"budget" + 0.011*"senat" + 0.010*"tower" + 0.009*"wiretap" + 0.007*"obama" + 0.007*"indic"
+ 0.006*"see" + 0.006*"claim" + 0.006*"thing"'),
 (6,
  u'0.016*"donald" + 0.015*"senat" + 0.012*"budget" + 0.009*"thing" + 0.009*"tower" + 0.008*"wife" + 0.008*"never" +
0.008*"committ" + 0.008*"wiretap" + 0.007*"us"'),
 (7,
  u'0.018*"budget" + 0.011*"senat" + 0.010*"wiretap" + 0.009*"donald" + 0.009*"tower" + 0.008*"obama" + 0.007*"commit
te" + 0.007*"us" + 0.006*"intel" + 0.006*"thing"'),
 (8,
  u'0.017*"budget" + 0.015*"senat" + 0.012*"tower" + 0.011*"donald" + 0.008*"wiretap" + 0.007*"obama" + 0.007*"commit
te" + 0.007*"us" + 0.007*"intel" + 0.006*"lie"'),
 (9,
  u'0.018*"budget" + 0.013*"tower" + 0.012*"donald" + 0.012*"senat" + 0.007*"wiretap" + 0.007*"ban" + 0.007*"knew" +
0.006*"thing" + 0.006*"obama" + 0.006*"indic"')]
```

PROJECT REPORT-SOCIAL MEDIA ANALYTICS

Following are the results obtained for 5 states -
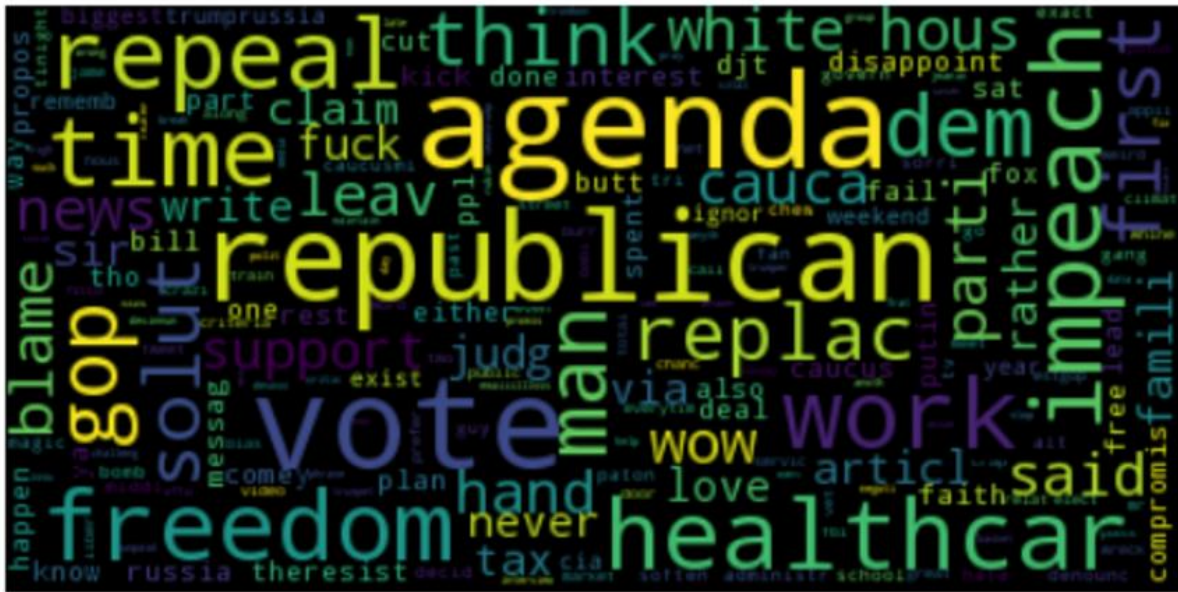
ANALYSIS FOR TEXAS STATE:

<u>Sentiment Analysis:</u>


```
Average of Polarity Scores: -0.0673165352853
Average of Subjectivity Scores: 0.331919374732
```

PROJECT REPORT-SOCIAL MEDIA ANALYTICS

Word Cloud:

Topic Modeling:

NMF

```
(1086, 2852)
Topic 0: trumprussia ryan presid obama peopl work think need support vote
Topic 1: know want care presid plan obama ryan new wrong donald
Topic 2: presid peopl news fake obama russia need white care golf
Topic 3: work ryan think watch support start presid care time lie
Topic 4: presid corrupt job list know bias law senat son question
Topic 5: pleas presid peopl work need agre els time think news
Topic 6: good presid peopl democrat republican parti form citizens polit noth
Topic 7: russia time peopl know day america tweet white plea resist
Topic 8: presid tock tick help think work need time right peopl
Topic 9: resign traitor unabl democraci unqualifi unfit honor constitut presid work
Topic 10: presid time need peopl know start resist america russia trumprussia
Topic 11: obama care new say job democrat oh tough president loser
Topic 12: corrupt new news vote elect chang america administr peopl explain
Topic 13: wrong want new thing donald presid love becom ryan left
Topic 14: white hous time russia peopl need plea open news told
Topic 15: compul lie actual everi weekend plus profits florida american total
Topic 16: trumprussia trump presid right new time block republican peopl need
Topic 17: tick tock presid time plea work peopl need think russia
Topic 18: tick tock time presid russia mr lose work trump need
Topic 19: evict work help va blind sweet rino teach presid plea
```
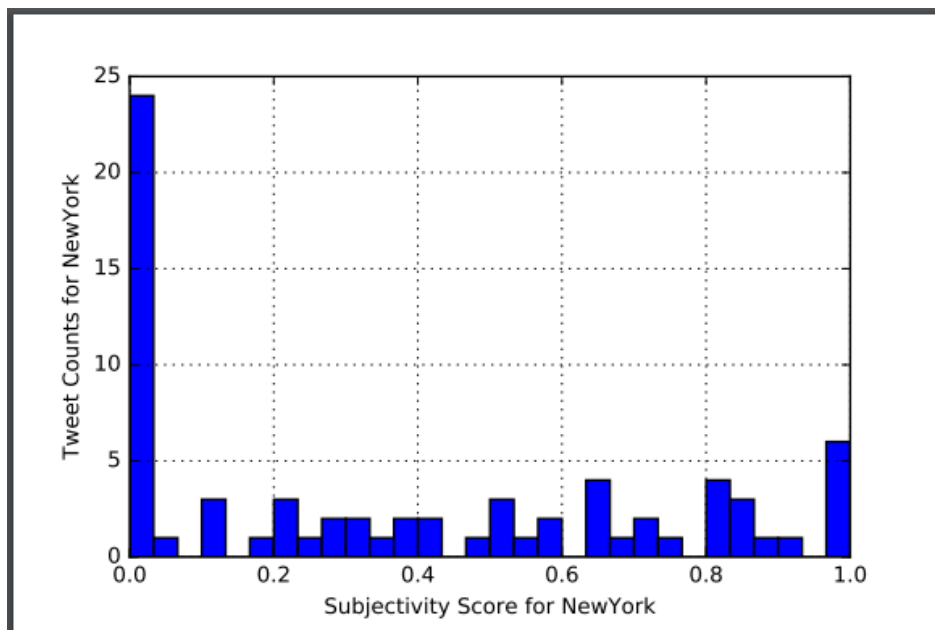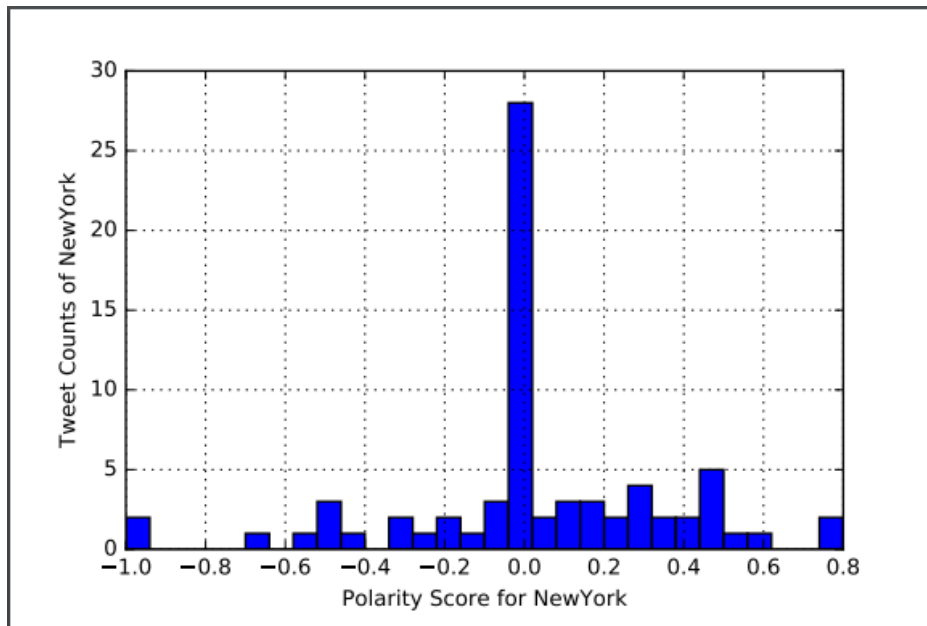
PROJECT REPORT-SOCIAL MEDIA ANALYTICS

LDA

```
[(0,
  u'0.011*"trumprussia" + 0.007*"get" + 0.007*"peopl" + 0.006*"ryan" + 0.006*"russia" + 0.006*"one" + 0.006*"need" + 0.006*"oba
ma" + 0.006*"care" + 0.005*"go"'),
 (1,
  u'0.008*"get" + 0.007*"trumprussia" + 0.007*"work" + 0.006*"time" + 0.006*"go" + 0.006*"need" + 0.006*"ryan" + 0.005*"care" +
0.005*"would" + 0.005*"obama"'),
 (2,
  u'0.010*"get" + 0.008*"ryan" + 0.007*"trumprussia" + 0.007*"peopl" + 0.006*"obama" + 0.006*"would" + 0.006*"one" + 0.005*"nee
d" + 0.005*"time" + 0.005*"work"'),
 (3,
  u'0.009*"get" + 0.009*"trumprussia" + 0.008*"obama" + 0.006*"peopl" + 0.006*"need" + 0.005*"go" + 0.005*"think" + 0.005*"woul
d" + 0.005*"time" + 0.005*"work"'),
 (4,
  u'0.008*"get" + 0.007*"would" + 0.007*"peopl" + 0.006*"need" + 0.006*"trumprussia" + 0.006*"think" + 0.006*"work" + 0.005*"ru
ssia" + 0.005*"news" + 0.005*"care"'),
 (5,
  u'0.008*"trumprussia" + 0.008*"peopl" + 0.007*"get" + 0.007*"time" + 0.006*"ryan" + 0.006*"russia" + 0.006*"would" + 0.006*"n
eed" + 0.006*"care" + 0.006*"one"'),
 (6,
  u'0.010*"get" + 0.009*"trumprussia" + 0.008*"would" + 0.007*"peopl" + 0.006*"work" + 0.005*"think" + 0.005*"us" + 0.005*"obam
a" + 0.005*"russia" + 0.005*"go"'),
 (7,
  u'0.010*"get" + 0.007*"peopl" + 0.006*"would" + 0.006*"work" + 0.005*"need" + 0.005*"ryan" + 0.005*"u" + 0.005*"go" + 0.005
*"care" + 0.005*"think"'),
 (8,
  u'0.009*"get" + 0.007*"go" + 0.006*"need" + 0.006*"trumprussia" + 0.006*"obama" + 0.005*"u" + 0.005*"would" + 0.005*"ryan" +
0.005*"news" + 0.005*"one"'),
 (9,
  u'0.011*"trumprussia" + 0.009*"get" + 0.008*"peopl" + 0.007*"time" + 0.006*"go" + 0.006*"need" + 0.006*"one" + 0.005*"ryan" +
0.005*"obama" + 0.005*"would"')]
```

PROJECT REPORT-SOCIAL MEDIA ANALYTICS

ANALYSIS FOR NEW YORK STATE:

Sentiment Analysis:

Average of Polarity Scores: 0.0364135151114

Average of Subjectivity scores: 0.378489675191

PROJECT REPORT-SOCIAL MEDIA ANALYTICS

Word Cloud:

Topic Modeling:

NMF

```
(1342, 3296)
Topic 0: presid donald know obama time ryan need hous think america
Topic 1: donald way hous thetrumpreport know say gop believ good white
Topic 2: russia day tax work congress hard lie say need el
Topic 3: presid fake weapon job peopl mission dscovr nasa strategi scienc
Topic 4: russia day run usa johnni immigr rotten anoth establishment like
Topic 5: nasa strategi dscovr scienc mission earth cut man evil approv
Topic 6: russia run presid hack planet geoengin experiments door open collus
Topic 7: els pleas agre presid donald peopl fake scienc mission trumprussia
Topic 8: tax hous white el someon leak use day cover season
Topic 9: approv program presid time lose rate lunch percent hillari obama
Topic 10: fake trumprussia new news job scienc obama mission ugh great
Topic 11: donald presid tri peopl know think need america putin time
Topic 12: hello ban saudi arabia september travel particip list vote el
Topic 13: donald thetrumpreport gop way say presid russia parti know fraud
Topic 14: hello ban saudi september arabia travel particip list ryan know
Topic 15: hello ban list particip travel september arabia saudi need donald
Topic 16: believ good men away scum rememb antifa fight trumpnat gir
Topic 17: balk maga trumprussia fake donald mission news mess approv tweet
Topic 18: presid gop outnumb support leak dem fool biggest confirm vice
Topic 19: presid trumprussia thought support tire right death putin day hit
```
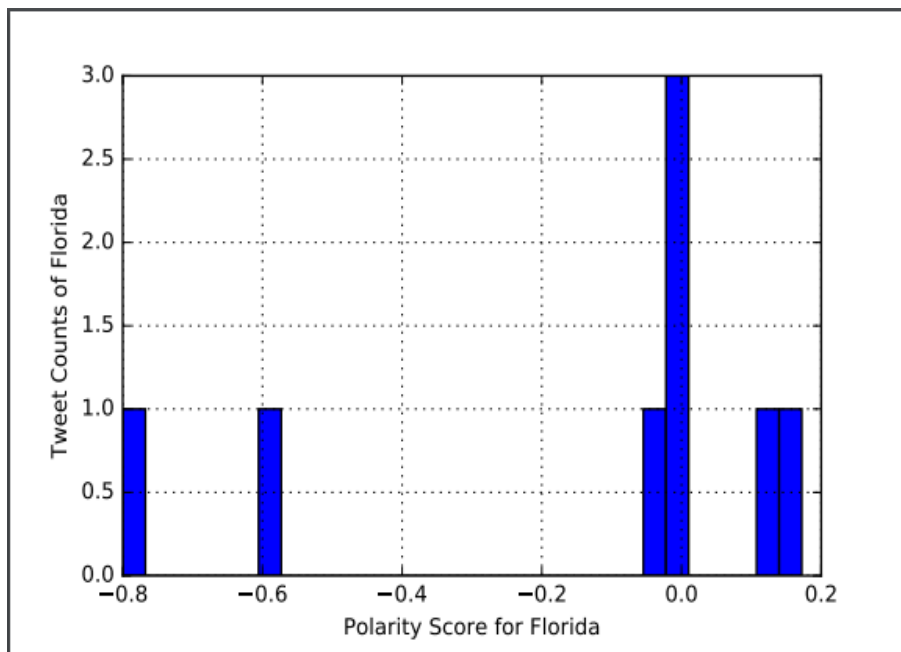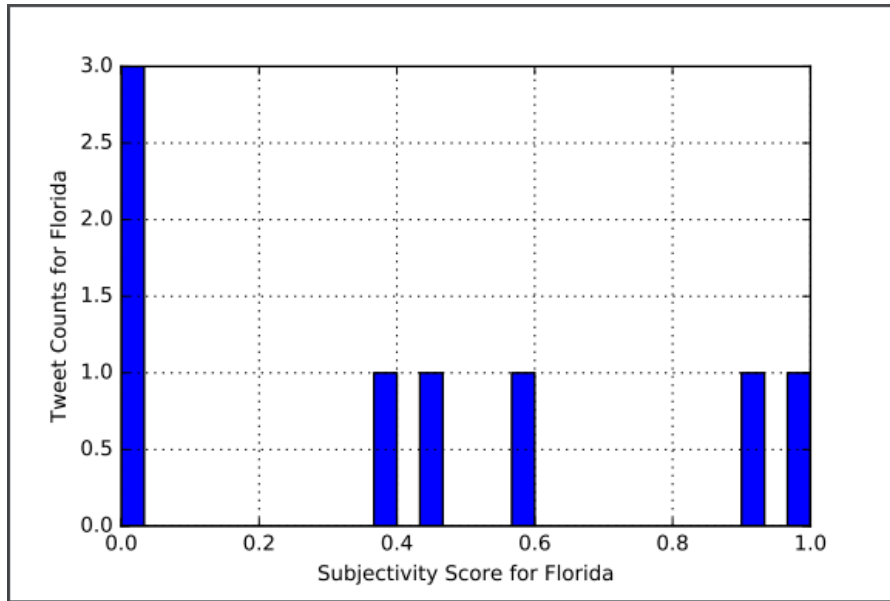
PROJECT REPORT-SOCIAL MEDIA ANALYTICS

LDA

```
[(0,
  u'0.009*"get" + 0.006*"go" + 0.006*"would" + 0.005*"obama" + 0.005*"see" + 0.005*"know" + 0.005*"one" + 0.005*"trumprussia" +
0.005*"think" + 0.005*"donald"'),
 (1,
  u'0.007*"go" + 0.007*"get" + 0.007*"donald" + 0.005*"need" + 0.005*"would" + 0.005*"obama" + 0.005*"america" + 0.005*"putin"
 + 0.005*"see" + 0.005*"ryan"'),
 (2,
  u'0.008*"go" + 0.008*"get" + 0.006*"donald" + 0.005*"know" + 0.005*"would" + 0.005*"time" + 0.004*"obama" + 0.004*"think" +
0.004*"trumprussia" + 0.004*"tax"'),
 (3,
  u'0.009*"get" + 0.007*"go" + 0.007*"would" + 0.006*"russia" + 0.006*"one" + 0.006*"need" + 0.005*"know" + 0.005*"donald" + 0.
004*"america" + 0.004*"think"'),
 (4,
  u'0.007*"go" + 0.006*"know" + 0.006*"get" + 0.005*"would" + 0.005*"donald" + 0.005*"trumprussia" + 0.005*"obama" + 0.005*"tim
e" + 0.004*"one" + 0.004*"putin"'),
 (5,
  u'0.010*"get" + 0.007*"go" + 0.006*"donald" + 0.005*"need" + 0.005*"would" + 0.005*"russia" + 0.005*"one" + 0.005*"know" + 0.
005*"hous" + 0.004*"obama"'),
 (6,
  u'0.008*"get" + 0.008*"go" + 0.007*"donald" + 0.006*"need" + 0.005*"would" + 0.005*"see" + 0.005*"peopl" + 0.004*"one" + 0.00
4*"know" + 0.004*"tax"'),
 (7,
  u'0.007*"get" + 0.007*"go" + 0.006*"would" + 0.005*"see" + 0.005*"trumprussia" + 0.005*"obama" + 0.005*"one" + 0.004*"suppor
t" + 0.004*"donald" + 0.004*"time"'),
 (8,
  u'0.008*"get" + 0.006*"would" + 0.006*"go" + 0.005*"donald" + 0.005*"tax" + 0.005*"ryan" + 0.005*"need" + 0.005*"know" + 0.00
5*"see" + 0.004*"one"'),
 (9,
  u'0.007*"get" + 0.007*"go" + 0.006*"know" + 0.006*"would" + 0.006*"donald" + 0.005*"support" + 0.005*"see" + 0.004*"hous" +
0.004*"need" + 0.004*"one"')]
```

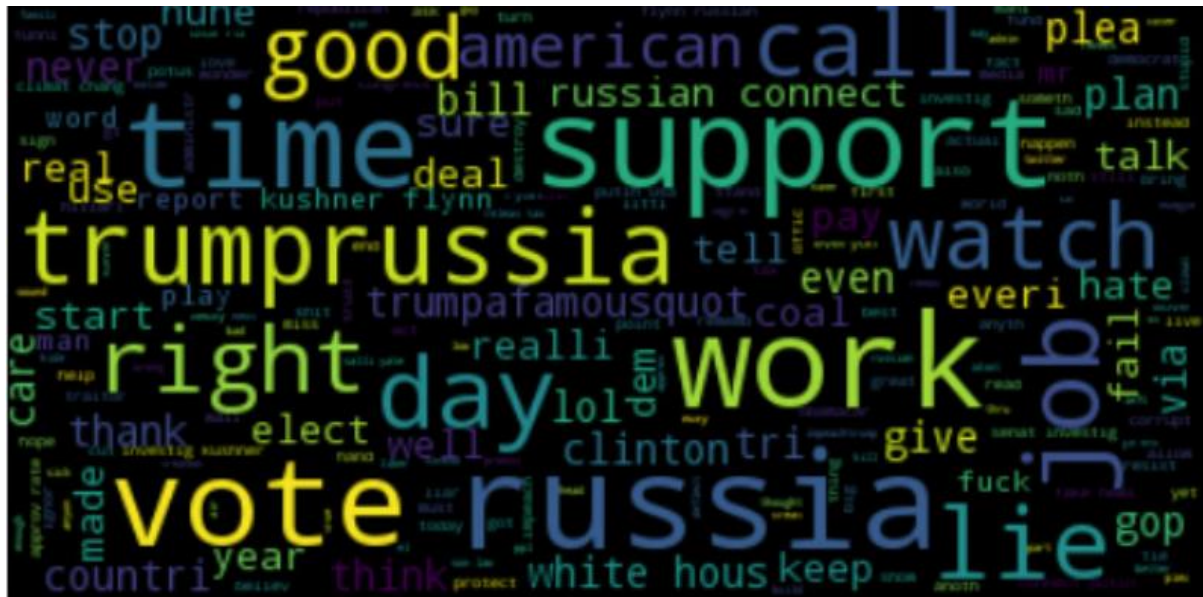**ANALYSIS FOR FLORIDA STATE:**

Sentiment Analysis:

PROJECT REPORT-SOCIAL MEDIA ANALYTICS

Average of Polarity Scores: -0.143034840593

Average of Subjectivity scores: 0.41810290404

Word Cloud:

PROJECT REPORT-SOCIAL MEDIA ANALYTICS

<u>Topic Modelling:</u>

<u>NMF</u>

```
Topic 0: presid russian obama work investig tax russia job kushner connect
Topic 1: russian investig kushner flynn senat putin connect ru usa wh
Topic 2: wh tax relea feb return broke record demand petit deadlin
Topic 3: news know presid obama fake support time thank donald trumprussia
Topic 4: caucus need freedom america presid dam donald poor nune lt
Topic 5: news know obama presid fake agr watch read god tax
Topic 6: allow sperbank russian hous kgb russia stop salli yate testifi
Topic 7: presid thank coal allah good sign trumpafamousquote america wake obama
Topic 8: ryancar obama presid approach need businessman scammi peopl fuck want
Topic 9: pleas releas agre presid obama committe russian promis stigmabas els
Topic 10: drug peopl wall kill matter young illeg cross important signed
Topic 11: america water need presid poor near nune maxin trumpafamousquote news
Topic 12: news presid know rep vote donald time agr white hous
Topic 13: obama presid russia testifi god reform undo block fact administr
Topic 14: putin tybe potushairdontcar downright trumpi work caucus ruin america georgia
Topic 15: obama presid donald trumpafamousquote know trumpafamousquot great day need support
Topic 16: news obama vote presid fake job stop away work russian
Topic 17: nune famili golf time know america spent limit vacations maralago
Topic 18: russian presid right sperbank son kushner work want russia time
Topic 19: donald obama golf presid know thank countri think support proud
```
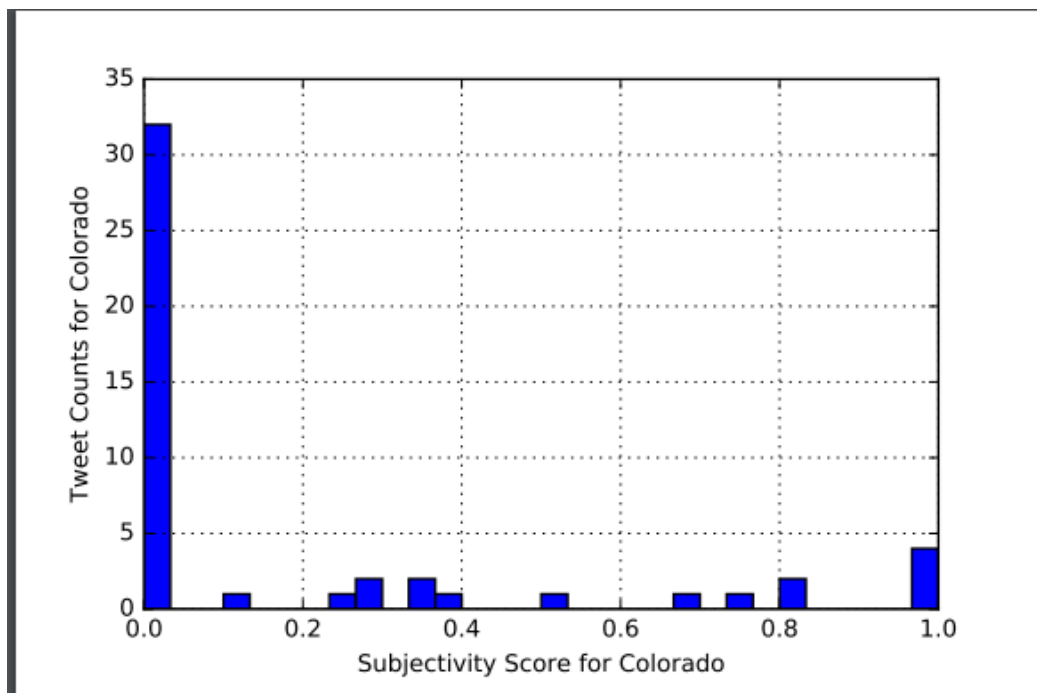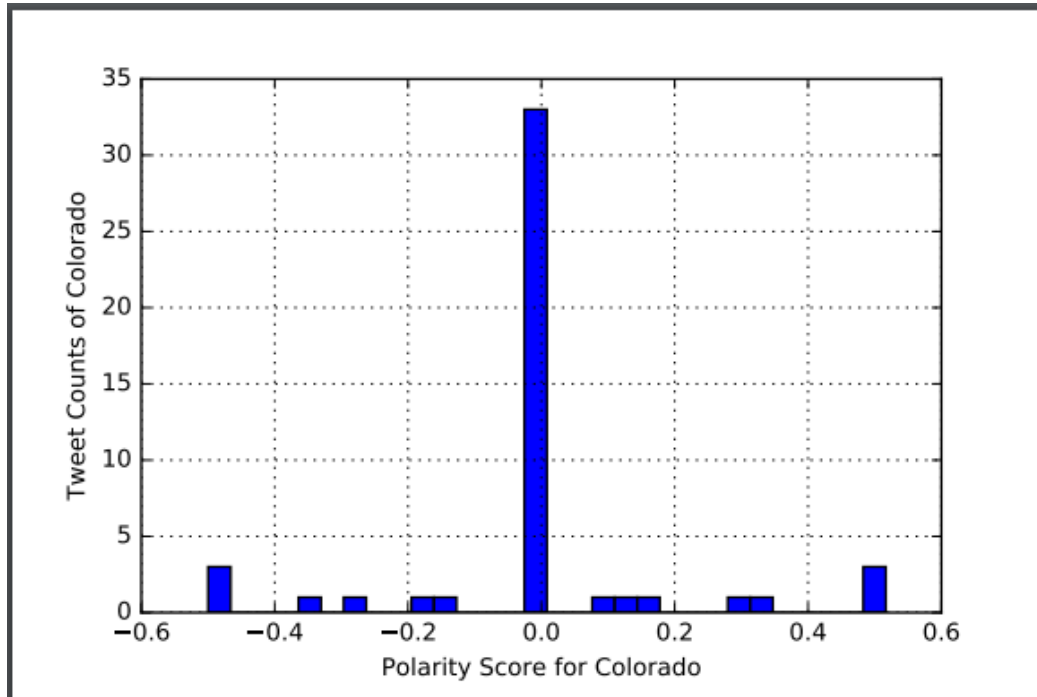
<u>**LDA**</u>

```
[(0,
  u'0.008*"get" + 0.006*"u" + 0.006*"obama" + 0.006*"russian" + 0.006*"support" + 0.005*"russia" + 0.005*"go" + 0.005
*"want" + 0.005*"work" + 0.005*"donald"'),
 (1,
  u'0.009*"obama" + 0.008*"u" + 0.007*"russian" + 0.006*"investig" + 0.006*"go" + 0.005*"tax" + 0.005*"work" + 0.005
*"support" + 0.005*"putin" + 0.004*"get"'),
 (2,
  u'0.008*"get" + 0.006*"go" + 0.006*"obama" + 0.006*"u" + 0.006*"russian" + 0.006*"one" + 0.004*"job" + 0.004*"russi
a" + 0.004*"tax" + 0.004*"know"'),
 (3,
  u'0.012*"russian" + 0.010*"investig" + 0.009*"kushner" + 0.009*"u" + 0.008*"wh" + 0.007*"putin" + 0.007*"connect" +
0.007*"usa" + 0.007*"tax" + 0.007*"flynn"'),
 (4,
  u'0.009*"u" + 0.008*"russian" + 0.008*"investig" + 0.007*"obama" + 0.006*"go" + 0.005*"one" + 0.005*"kushner" + 0.0
05*"putin" + 0.005*"work" + 0.004*"us"'),
 (5,
  u'0.008*"u" + 0.006*"go" + 0.005*"obama" + 0.005*"russian" + 0.005*"investig" + 0.005*"support" + 0.005*"get" + 0.0
04*"need" + 0.004*"putin" + 0.004*"russia"'),
 (6,
  u'0.009*"russian" + 0.008*"u" + 0.007*"obama" + 0.006*"tax" + 0.005*"investig" + 0.005*"get" + 0.005*"take" + 0.005
*"work" + 0.005*"go" + 0.005*"russia"'),
 (7,
  u'0.008*"obama" + 0.007*"u" + 0.006*"get" + 0.006*"go" + 0.005*"russian" + 0.004*"job" + 0.004*"time" + 0.004*"tax"
+ 0.004*"investig" + 0.004*"want"'),
 (8,
  u'0.008*"u" + 0.006*"get" + 0.006*"obama" + 0.005*"take" + 0.005*"russian" + 0.005*"russia" + 0.005*"go" + 0.005*"j
ob" + 0.004*"support" + 0.004*"putin"'),
 (9,
  u'0.008*"u" + 0.008*"obama" + 0.007*"go" + 0.006*"get" + 0.005*"us" + 0.005*"work" + 0.005*"need" + 0.005*"one" +
0.005*"job" + 0.005*"support"')]
```

PROJECT REPORT-SOCIAL MEDIA ANALYTICS

ANALYSIS FOR COLORADO STATE:
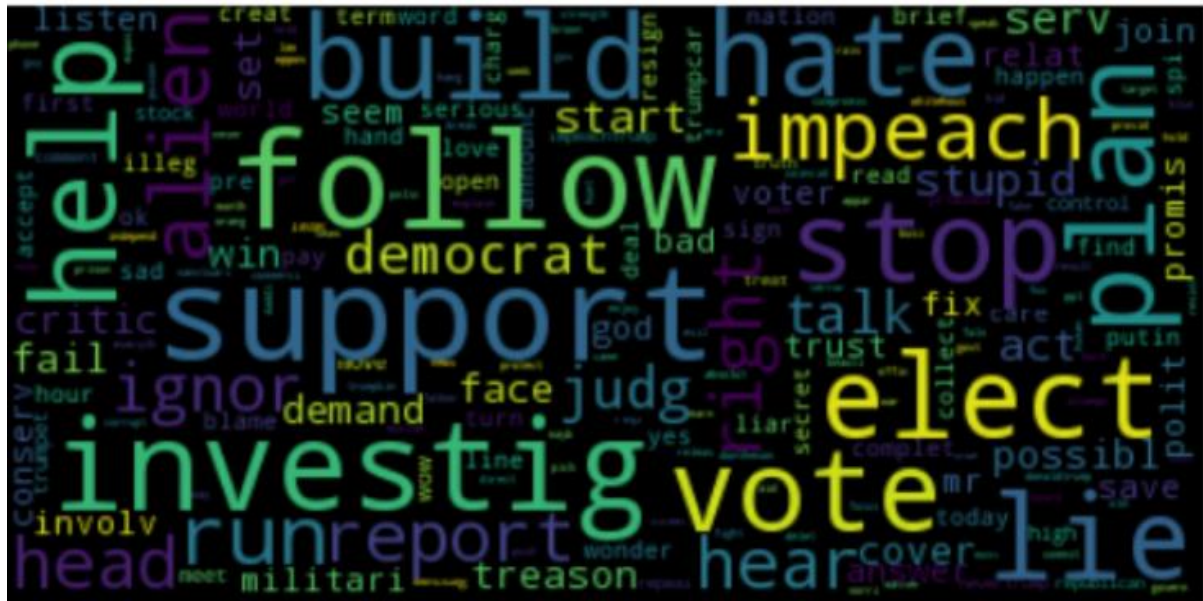
Sentiment Analysis:

Average of Polarity Scores: 0.0017898478836

Average of Subjectivity Scores: 0.198916997354

PROJECT REPORT-SOCIAL MEDIA ANALYTICS

Word Cloud:

Topic Modeling:

```
Topic 0: puppet treasonous need russian spec impeachtrumpnow independ prosecutor comm trumprussia
Topic 1: america thing reek ingratitud incompet grift offic need putin american
Topic 2: biggest work tweet demand protest prosecutor spec impeachtrumpnow thank awesom
Topic 3: zero need love stupid america want got peopl freedom support
Topic 4: trumprussia rantfortoday gop today need know trump ccot resist traitor
Topic 5: profess want treat inte payback sweet trumprussia tap emaiiilllsss raid
Topic 6: putin thing american sad funni think ppl tell work care
Topic 7: america thing reek incompet ingratitud grift offic potustrump busi need
Topic 8: wake theresistance impeachtrumpnow american russian promis trumprussia rantfortoday bindtrump pleas
Topic 9: becom obama blame prez baby sure man boss need american
Topic 10: need canadian america russia zero famili come yes new repeal
Topic 11: donald way kid obama thing time th blame good tell
Topic 12: stupid spi bush jeb obama accept got kid donald way
Topic 13: come know let shade office inaugur darker treasury raid emaiiilllsss
Topic 14: need vote offic news busi american care refug inform pelosi
Topic 15: need support stupid want got america mean obama good accept
Topic 16: protest paid yup um credibility lose work hous women white
Topic 17: probabl emaiiilllsss raid treasury profess inte payback sweet treat want
Topic 18: tweet donald presid open secret messag realli bomb bath magical
Topic 19: rantfortoday ccot free know time govt need control resist mr
```
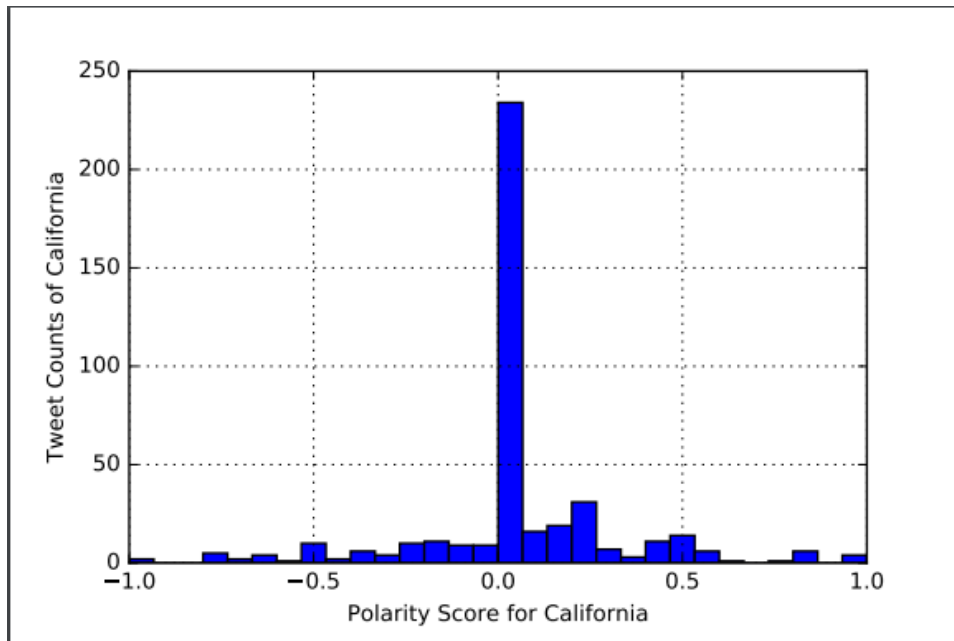
PROJECT REPORT-SOCIAL MEDIA ANALYTICS
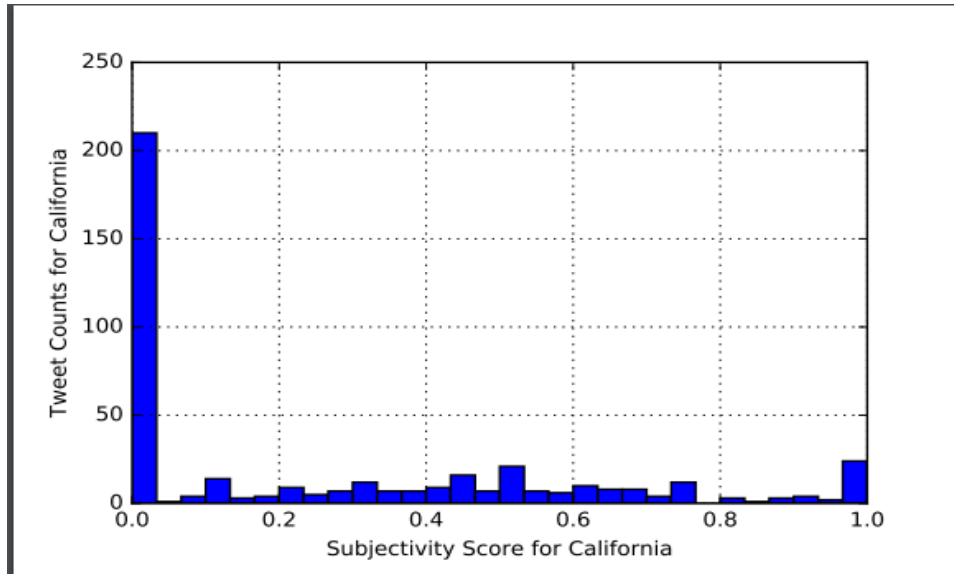
<u>LDA</u>

```
[(0,
  u'0.068*"spec" + 0.052*"need" + 0.051*"impeachtrumpnow" + 0.047*"comm" + 0.046*"treasonous" + 0.044*"russian" + 0.0
41*"independ" + 0.035*"puppet" + 0.032*"prosecutor" + 0.006*"us"'),
 (1,
  u'0.052*"russian" + 0.049*"treasonous" + 0.048*"impeachtrumpnow" + 0.048*"independ" + 0.048*"comm" + 0.044*"prosecu
tor" + 0.043*"need" + 0.041*"spec" + 0.036*"puppet" + 0.006*"trumprussia"'),
 (2,
  u'0.064*"need" + 0.053*"independ" + 0.053*"prosecutor" + 0.053*"impeachtrumpnow" + 0.052*"treasonous" + 0.049*"spe
c" + 0.042*"puppet" + 0.042*"russian" + 0.041*"comm" + 0.004*"us"'),
 (3,
  u'0.066*"independ" + 0.065*"puppet" + 0.059*"need" + 0.058*"treasonous" + 0.053*"comm" + 0.053*"prosecutor" + 0.044
*"spec" + 0.041*"impeachtrumpnow" + 0.039*"russian" + 0.006*"rantfortoday"'),
 (4,
  u'0.056*"treasonous" + 0.053*"puppet" + 0.047*"prosecutor" + 0.045*"impeachtrumpnow" + 0.043*"spec" + 0.042*"comm"
+ 0.039*"russian" + 0.034*"independ" + 0.030*"need" + 0.005*"trumprussia"'),
 (5,
  u'0.052*"russian" + 0.051*"impeachtrumpnow" + 0.048*"prosecutor" + 0.044*"independ" + 0.042*"puppet" + 0.041*"spec"
+ 0.038*"comm" + 0.038*"need" + 0.033*"treasonous" + 0.006*"trumprussia"'),
 (6,
  u'0.047*"need" + 0.042*"spec" + 0.042*"impeachtrumpnow" + 0.041*"comm" + 0.041*"prosecutor" + 0.040*"independ" + 0.
037*"puppet" + 0.036*"russian" + 0.035*"treasonous" + 0.005*"trumprussia"'),
 (7,
  u'0.065*"russian" + 0.060*"comm" + 0.058*"spec" + 0.058*"puppet" + 0.057*"impeachtrumpnow" + 0.055*"prosecutor" +
0.054*"treasonous" + 0.054*"need" + 0.050*"independ" + 0.005*"rantfortoday"'),
 (8,
  u'0.061*"need" + 0.058*"russian" + 0.053*"independ" + 0.052*"prosecutor" + 0.051*"impeachtrumpnow" + 0.047*"spec" +
0.046*"puppet" + 0.042*"comm" + 0.036*"treasonous" + 0.005*"us"'),
 (9,
  u'0.036*"need" + 0.034*"prosecutor" + 0.033*"impeachtrumpnow" + 0.032*"independ" + 0.032*"comm" + 0.030*"spec" + 0.
029*"russian" + 0.028*"treasonous" + 0.025*"puppet" + 0.006*"america"')]
```

<u>ANALYSIS FOR CALIFORNIA STATE:</u>

<u>Sentiment Analysis:</u>

PROJECT REPORT-SOCIAL MEDIA ANALYTICS

Average of Polarity Scores: 0.0342364656885

Average of Subjectivity Scores: 0.263082211467

Word Cloud:

Topic Modelling:

NMF:

PROJECT REPORT-SOCIAL MEDIA ANALYTICS

```
Topic 0: russia presid donald work want think america good american day
Topic 1: russia day lie golf hous republican donald loser presid pro
Topic 2: pleas hear trumprussia dea russian public intel cours tomorrow cmmte
Topic 3: dea peopl isi mil beat vote plea president cnn ilwaco
Topic 4: want cbs good donald america think peopl golf work dc
Topic 5: newsweek dea trumprussia hear isi pleas real peopl russian neva
Topic 6: peopl know vote big comment buy yeah enemi mirror dowd
Topic 7: dea plea russian cour newsweek ilwaco promi vote president report
Topic 8: lie hous day care peopl pro russia complic someth trump
Topic 9: bbc pleas dea react isi trumprussia chris isis presid peopl
Topic 10: continu huge legisl confus followers amjoy waffl countri believ live
Topic 11: shall ur trumpcar love think day care gorsuch anoth goe
Topic 12: hear dea pleas intel public russian trumprussia tomorrow cmmte cours
Topic 13: good american think day time putin hillari news russia use
Topic 14: russia donald good time think golf gov real hate tweet
Topic 15: hear trumprussia intel public tomorrow cmmte russian tie open right
Topic 16: newsweek isi trumprussia dea real pleas neva russia peopl russian
Topic 17: say vote trumprussia care run come gorsuch turn smell test
Topic 18: russia donald peopl hous senat prison pro presid putin women
Topic 19: dea support vote cour russia saw gonna promi want liar
```

LDA

```
[(0,
  u'0.006*"get" + 0.006*"trumprussia" + 0.005*"work" + 0.005*"go" + 0.005*"donald" + 0.005*"support" + 0.005*"think" + 0.005*"p
eopl" + 0.005*"good" + 0.005*"need"'),
 (1,
  u'0.007*"russia" + 0.006*"work" + 0.006*"see" + 0.006*"think" + 0.006*"ryan" + 0.005*"peopl" + 0.005*"get" + 0.005*"want" +
 0.005*"america" + 0.005*"support"'),
 (2,
  u'0.008*"russia" + 0.007*"go" + 0.006*"donald" + 0.006*"get" + 0.005*"think" + 0.005*"hous" + 0.005*"want" + 0.005*"american"
 + 0.005*"america" + 0.005*"one"'),
 (3,
  u'0.007*"trumprussia" + 0.006*"russia" + 0.006*"american" + 0.005*"peopl" + 0.005*"support" + 0.005*"one" + 0.005*"work" + 0.
005*"dea" + 0.005*"get" + 0.005*"russian"'),
 (4,
  u'0.008*"russia" + 0.007*"get" + 0.006*"golf" + 0.006*"go" + 0.006*"peopl" + 0.005*"via" + 0.005*"donald" + 0.005*"day" + 0.0
05*"good" + 0.005*"want"'),
 (5,
  u'0.008*"russia" + 0.006*"get" + 0.006*"via" + 0.006*"donald" + 0.006*"america" + 0.005*"peopl" + 0.005*"russian" + 0.005*"wa
nt" + 0.005*"ryan" + 0.004*"one"'),
 (6,
  u'0.006*"peopl" + 0.006*"via" + 0.006*"get" + 0.006*"russia" + 0.005*"support" + 0.005*"would" + 0.005*"isi" + 0.005*"need" +
0.005*"see" + 0.005*"call"'),
 (7,
  u'0.007*"trumprussia" + 0.006*"get" + 0.006*"see" + 0.006*"peopl" + 0.006*"think" + 0.006*"via" + 0.005*"russia" + 0.005*"kno
w" + 0.005*"u" + 0.005*"russian"'),
 (8,
  u'0.008*"go" + 0.007*"russia" + 0.006*"get" + 0.006*"peopl" + 0.006*"trumprussia" + 0.005*"donald" + 0.005*"think" + 0.005*"s
ee" + 0.005*"golf" + 0.005*"know"'),
 (9,
  u'0.008*"russia" + 0.007*"peopl" + 0.006*"trumprussia" + 0.006*"good" + 0.006*"want" + 0.006*"via" + 0.006*"go" + 0.006*"kno
w" + 0.005*"russian" + 0.005*"day"')]
```

PROJECT REPORT-SOCIAL MEDIA ANALYTICS

Insights:

As we understand, it is easy for people to express their opinion, emotions, ideas, or frustration over social media platform. With the above results through analysis, we have derived following insights.

From the 10k tweets collected, we can derive that Trump as a President has gained lot of attention during the elections because of his policies like Muslim ban, travel ban, budget cuts and proposals. The sentiment score suggests that sentiments are neutral because there are positive comments by his supporters as well as negative comments from haters or disappointed citizens.

After he became the President, his connections with Russia and wiretapping made a lot of news due to which he gained negative tweets on social media. The travel ban had mixed reactions from the citizens where the supporters were happy that their President Elect lived up to the promises made during the elections. This led to good number of positive tweets. Hence our overall observations along with sentiment analysis, we think Trump's influence in United States is neutral at the moment.

Some of our interpretations based on the results of word clouds, histograms, and topics of 5 states, following are the observations:

- 'trumprussia' is the most common topics between the 5 states currently.
- 'ISI' is the unique topic among 5 states. We could interpret that some people tweeting from California are vocal about terrorist groups 'ISI'.
- Colorado state was observed to be quite neutral. While some spoke more about impeachment of Trump on basis of treachery and accusations of Trump's connection with Russia, some of them are positive because of words 'trumpcare', 'help', 'support'.
- Contrary to the ideal assumption that a blue state would be against Trump and red state for Trump, we observed New York is the **most positive state** and Texas is **most negative state** based on polarity scores.
- Although average of polarity score is positive for New York state, people were disappointed about Trump's proposal for cut in NASA's Earth Science program.
- California and New York have many common topics like 'Russia,', 'lie', 'wall'. This can be explained as these two states are democratic states so they are talking about few common topics.

PROJECT REPORT-SOCIAL MEDIA ANALYTICS

References:

1. Video lectures, code by Prof Gene Moo Lee
2. https://stackoverflow.com
3. https://en.wikipedia.org/wiki/Red_states_and_blue_states
4. https://regex101.com
5. http://www.latimes.com
6. https://www.google.com/maps