



A3: Business Insight Report : Sentiment Analysis : Speeches by Donald Trump and Bernie Sanders

Jaisanker Venugopalan Nair

Text Analytics - DAT- 5317 – FMSBA3

Thomas Kurnicki

February 14, 2020

Background

With the 2020 US presidential election campaign on full throttle, all the major candidates are traveling around the country to win the primary elections/caucuses and gain the required delegates to get the nomination from their respective parties. At the time of this article, president incumbent of the USA, Donald J. Trump, is the clear front runner for the Republican nomination and enjoys a loyal base of right-wing voters mostly from the central and western United States. On the other side, among the eight candidates running for democratic nomination is the independent senator from Vermont, Bernie Sanders, campaigning on the platform of democratic socialism, an ideology to which the voters of the USA are still warming up. His voter base consists of mostly young professionals and middle class in the more urbanized areas of the United States.

Analysis

For this article, speeches of both candidates in the last four years were collected and analyzed for similarities, uniqueness, and trends of their respective campaigns.

The frequency of words used by the candidates shows alignment to the platforms each of them is campaigning. President Trump mostly engages in speeches that incite nationalist feelings, tough immigration reforms including building a wall along the southern border of US with Mexico, tax cuts, his deal-making capabilities, more stringent trade agreements with other countries like China.

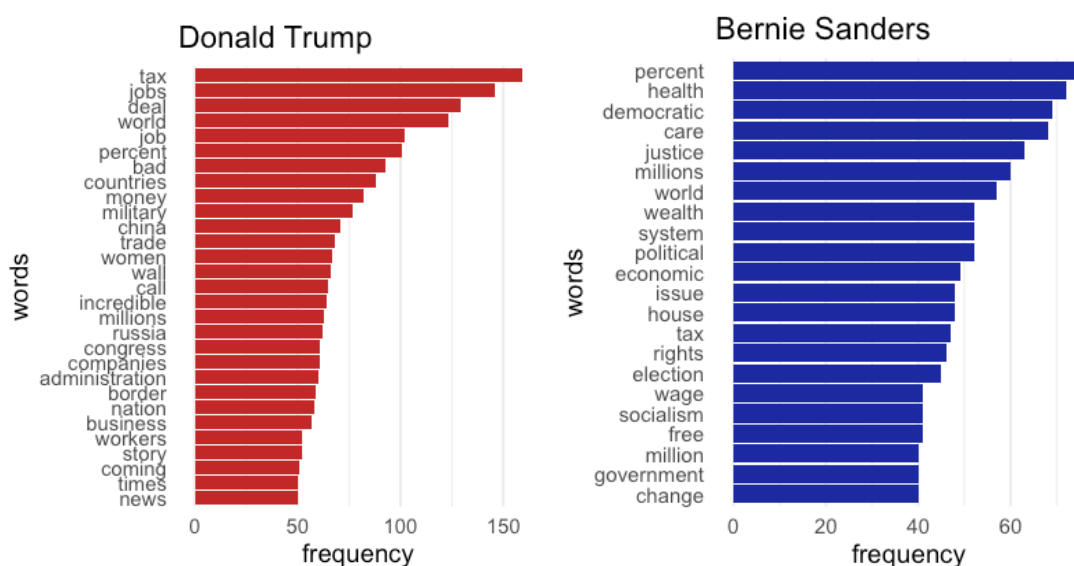


Fig.1 : Frequency of tokens

Senator Sanders talks about taxing the top 1% of wealthy individuals, social/political/economic reforms to help poorer sections of the community, healthcare reforms, free college education, and similar issues.

It is interesting to note that though talking about different subjects to different audience, both the candidates chose to stick to negative sentiments (AFINN score - Trump: -167, Bernie: -190) (Bing Score – Trump – -555/400, Bernie: -371/218) and surprise to gain the attention of

their respective bases and subsequently gain their trust as is evident from the sentiment word cloud below.



Fig.2 : Sentiment Word Cloud

On creating bigrams for speeches of each candidate, distinct characters of each campaign start to appear. As an example, one of the most used bigrams in President Trump's speeches is fake news, a term commonly used by the president to discredit the outlets that report news which the president does not necessarily approve. Similarly, another frequently occurring bigram – Paris accord, which the United States Pulled out of, refers to Trump's claim of climate change and global warming being a hoax perpetrated by China.

On the other hand, one of the most occurring bigrams in Senator Sanders' speeches are health care, referring to his campaign promise of making health care a human right. Similarly, another

bigram, Wallstreet, refers to his position on large banks being too big to fail and breaking up these large banks into smaller banks.

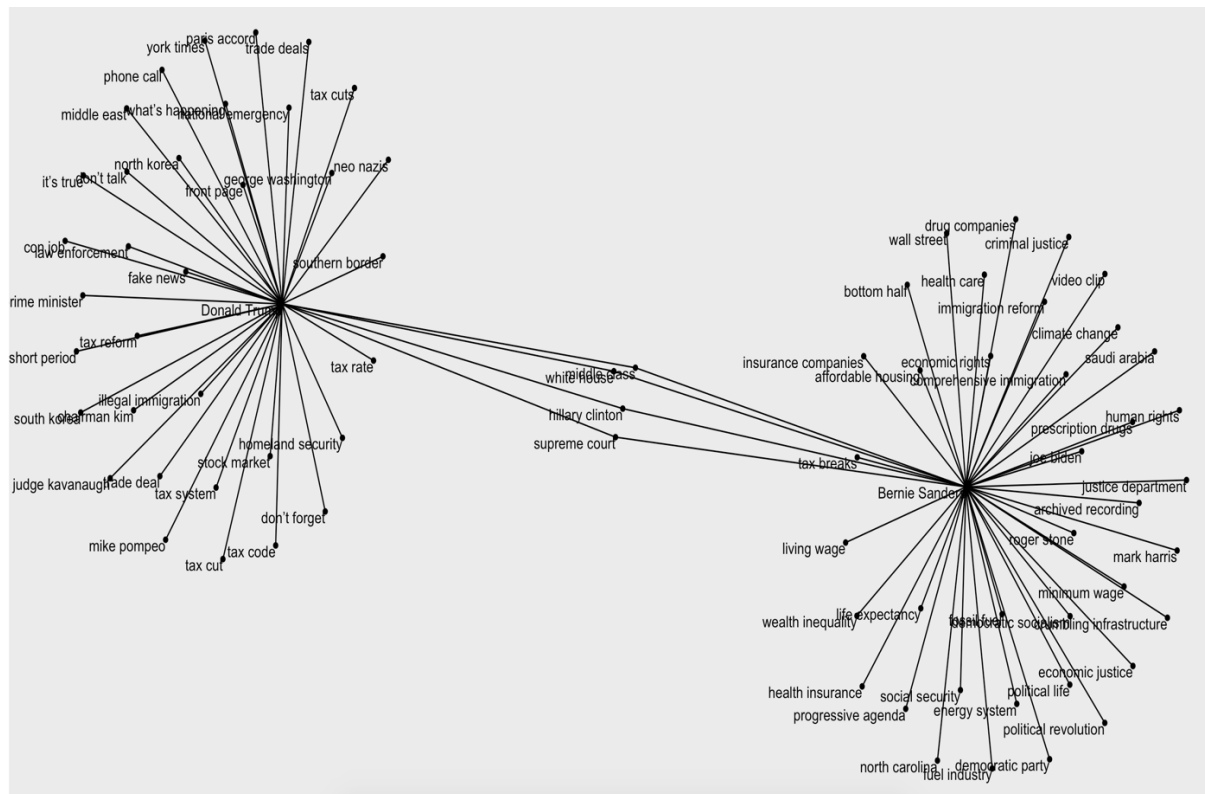


Fig.3 : Sentiment Word Cloud

The bigram graph shows the polarizing nature of issues raised by each candidate and correspondingly the type of vote bases each of them is trying to appeal. However, it is worth noting that, along with the differences, few topics are common among both. These topics can be of utmost importance for each while trying to woo voters in swing states and strongholds of the other candidate.

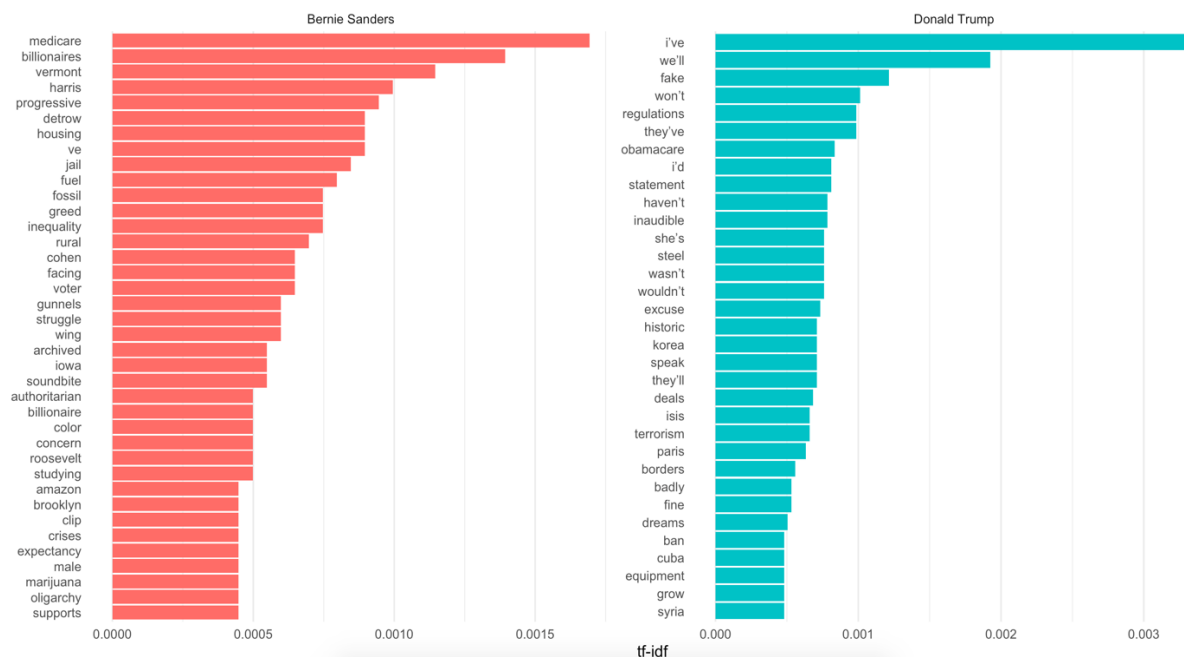


Fig.4 : TF-IDF

The TF IDF analysis reveals the prominent campaign characteristics of each candidate. One of the major promises that helped President Trump win his 2016 elections was repealing Obamacare and replacing it with a better one. Other promises, such as withdrawing from Syria, obliteration of ISIS, banning immigrants, et cetera, also shows up on analysis as having high significance.

As for Senator Sanders, TF-IDF shows high significance for Medicare, along with his progressive agenda and the crusade against billionaires and is consistent with his campaign promises.

Recommendations

Senator Sanders has a very deep and narrow base (Time,2019). While the progressive wing of the democratic party is on board with Sen Sanders', it is necessary for senator sanders to gain a broader vote base get nominated in the primaries and to defeat President Trump in the upcoming presidential elections. As President Trump's loyal base of conservative Republican voters is not likely to switch votes in the presidential elections, it leaves the Sanders campaign with two potential groups of voters : -

- Moderate Democrats: who consider Bernie Sanders to be too progressive and holds the key to his nomination as the democratic candidate for the election in the primaries and the presidential election.
- Moderate Republicans: who voted for Donald Trump during 2016 and consider him to be not worthy of a second term. It is worth noting that they will be eligible to vote for a republican or a democratic nominee in the presidential election.

For group (i), making recommendations will be outside the scope of this article as the analysis did not include other democratic candidates for primaries and is open for another study. The recommendation of this article will be directed towards appealing to group (ii), i.e., Moderate republicans.

Sanders Campaign should be trying to connect to the Moderate Republican voters (who aligned with Health care issues and middle-class issues raised by both candidates). Taking into consideration that the TF-IDF analysis of both candidates showed significant keywords aligned with the identity of their campaigns and Donald Trump went on to win the elections, Sanders Campaign should research the significant campaign promises made by Donald Trump and how it resonated with moderate republicans and whether they voted for him.

Secondly, though the bigram graphs show a highly polarized selection of topics between the two campaigns, it also shows a few common issues, namely Health care and Middle class, were among them. Considering that these topics were important for leaders and voters on either side of the aisle, efforts must be made to understand the needs of the moderate republicans. On receiving inputs from these efforts, the Sanders Campaign can reach out to the group through aggressive targeted advertisements to show the failure of Trump's presidency to enact a better health care act in place of Obamacare. The same strategy can be applied to expose President Trump's Tax initiatives that resulted in a lot of middle-class families having to pay more taxes (Bloomberg, 2018) as opposed to his campaign promises.

Finally, further analyses should be commissioned to collect more data about other Democratic candidates and the moderate democrats(group – ii) to understand the sentiments better and help policy formation.To conclude, the analysis showed plenty of similarities in the way both

candidates were trying to appeal to their bases and, at the same time, highlighted the uniqueness of each candidates' ideas. Bigram analysis sheds light on specific keywords each candidate used during their speeches. TF-IDF showed the significance of these bigrams in their addresses. Finally, for recommendations, the analysis showed the voter groups who have to target and how to achieve that goal.

Appendix I - References

1. Bernie Sanders on his 2020 run. TRANSCRIPT: 02/21/2019, All In w. Chris Hayes. (2019, February 21). Retrieved from <http://www.msnbc.com/transcripts/all-in/2019-02-21>
2. Board, E. (2019, November 20). Opinion | The global reach of Trump's 'fake news' outrage. Retrieved from https://www.washingtonpost.com/opinions/global-opinions/trump-is-spreading-his-fake-news-rhetoric-around-the-world-thats-dangerous/2019/11/19/a7b0a4c6-0af5-11ea-97ac-a7ccc8dd1ebc_story.html
3. Choi, D. (2019, March 8). Read Bernie Sanders' full speech from his first Iowa 2020 campaign rally. Retrieved from <https://www.businessinsider.com/bernie-sanders-iowa-2020-rally-speech-transcript-2019-3>
4. CNN Hosts Sen. Bernie Sanders (I-VT) Presidential Candidate in a Town Hall Discussion. Aired 9-10p ET. (2019, April 22). Retrieved February 13, 2020, from <http://edition.cnn.com/TRANSCRIPTS/1904/22/se.03.html>
5. Cohn, N. (2015, July 30). Support for Bernie Sanders Is Deep but Narrow. Retrieved from <https://www.nytimes.com/2015/07/31/upshot/support-for-bernie-sanders-is-deep-but-narrow.html>
6. Denchak, M. (2019, December 2). Paris Climate Agreement: Everything You Need to Know. Retrieved from <https://www.nrdc.org/stories/paris-climate-agreement-everything-you-need-know>
7. Detrow, S. (2019, December 16). 'The Speech' That Made Bernie Sanders A National Figure. Retrieved from <https://www.npr.org/2019/12/16/788597876/the-speech-that-made-bernie-sanders-a-national-figure>
8. Donald J. Trump for President. (n.d.). Retrieved from <https://www.donaldjtrump.com/about>
9. Donald Trump Speech in Miami, Sept. 16 • Transcripts . FactCheck.org. (2017, January 30). Retrieved from <https://transcripts.factcheck.org/donald-trump-speech-in-miami-sept-16/>
10. Full Bernie Sanders Speech on Economic Justice, Healthcare, Opposing Trump & Ending the War in Yemen. (n.d.). Retrieved from https://www.democracynow.org/2018/11/30/full_bernie_sanders_speech_on_economic
11. Full transcript: Sen. Bernie Sanders on "Face the Nation". (2019, June 23). Retrieved from <https://www.cbsnews.com/news/full-transcript-sen-bernie-sanders-on-face-the-nation/>
12. Golshan, T. (2019, June 12). Read: Bernie Sanders defines his vision for democratic socialism in the United States. Retrieved from <https://www.vox.com/2019/6/12/18663217/bernie-sanders-democratic-socialism-speech-transcript>
13. Kozlowska, H. (2017, May 30). Here's Bernie Sanders' full commencement speech at Brooklyn College. Retrieved from <https://qz.com/994609/transcript-heres-bernie-sanders-full-commencement-speech-at-brooklyn-college/>
14. Linn, J. C. F., Steverman, B., & Merrill, D. (2018, December 18). A Year After the Middle Class Tax Cut, the Rich Are Winning. Retrieved February 14, 2020, from <https://www.bloomberg.com/graphics/2018-tax-plan-consequences/>
15. Medicare for All. (n.d.). Retrieved from <https://berniesanders.com/issues/medicare-for-all/>

16. Memo: Polling Shows Bernie Sanders Builds Strong Base to Defeat Trump. (n.d.). Retrieved from <https://berniesanders.com/en/press/memo-polling-shows-bernie-sanders-builds-strong-base-defeat-trump/>
17. President Donald J. Trump Inaugural Address • Transcripts . FactCheck.org. (2017, January 25). Retrieved from <https://transcripts.factcheck.org/president-donald-j-trump-inaugural-address/>
18. President Trump on the Paris Climate Accord • Transcripts . FactCheck.org. (2017, June 3). Retrieved from <https://transcripts.factcheck.org/president-trump-paris-climate-accord/>
19. President Trump Press Conference • Transcripts . FactCheck.org. (2017, February 21). Retrieved from <https://transcripts.factcheck.org/president-trump-press-conference/>
20. President Trump Statement on Immigration • Transcripts . FactCheck.org. (2017, January 30). Retrieved from <https://transcripts.factcheck.org/president-trump-statement-immigration/>
21. President Trump's Address to a Joint Session of Congress • Transcripts . FactCheck.org. (2017, March 1). Retrieved from <https://transcripts.factcheck.org/president-trumps-address-joint-session-congress/>
22. President Trump's National Emergency Remarks • Transcripts . FactCheck.org. (2019, February 18). Retrieved from <https://transcripts.factcheck.org/382-2/>
23. President Trump's U.N. Press Conference • Transcripts . FactCheck.org. (2018, September 28). Retrieved from <https://transcripts.factcheck.org/president-trumps-u-n-press-conference/>
24. Real Wall Street Reform. (n.d.). Retrieved from <https://berniesanders.com/issues/real-wall-street-reform/>
25. Remarks by President Trump and Vice President Pence at CIA Headquarters • Transcripts . FactCheck.org. (2017, January 25). Retrieved from <https://transcripts.factcheck.org/remarks-president-trump-vice-president-pence-cia-headquarters/>
26. Remarks by President Trump at Tax Reform Event • Transcripts . FactCheck.org. (2017, September 28). Retrieved from <https://transcripts.factcheck.org/remarks-president-trump-tax-reform-event/>
27. Remarks by President Trump at the World Economic Forum • Transcripts . FactCheck.org. (2018, January 26). Retrieved from <https://transcripts.factcheck.org/remarks-president-trump-world-economic-forum/>
28. Remarks by President Trump on Infrastructure, Charlottesville • Transcripts . FactCheck.org. (2017, August 18). Retrieved from <https://transcripts.factcheck.org/remarks-president-trump-infrastructure-charlottesville/>
29. Remarks by President Trump on Iran • Transcripts . FactCheck.org. (2020, January 9). Retrieved from <https://transcripts.factcheck.org/remarks-by-president-trump-on-iran/>
30. Remarks by President Trump on Tax Reform • Transcripts . FactCheck.org. (2017, September 5). Retrieved from <https://transcripts.factcheck.org/remarks-president-trump-tax-reform/>
31. Remarks by the President at the NRCC Dinner • Transcripts . FactCheck.org. (2017, March 22). Retrieved from <https://transcripts.factcheck.org/remarks-president-nrcc-dinner/>

32. Remarks by the President in Nashville, Tennessee • Transcripts . FactCheck.org. (2017, March 22). Retrieved from <https://transcripts.factcheck.org/remarks-president-nashville-tennessee/>
33. Sanders Response to 2019 State of the Union Address. (2019, February 5). Retrieved from <https://www.sanders.senate.gov/newsroom/press-releases/sanders-response-to-2019-state-of-the-union-address>
34. Staff, W. P. (2016, July 26). Transcript: Bernie Sanders's full speech at the 2016 DNC. Retrieved from <https://www.washingtonpost.com/news/post-politics/wp/2016/07/26/transcript-bernie-sanderss-full-speech-at-the-2016-dnc/>
35. State of the Union Address • Transcripts . FactCheck.org. (2019, February 7). Retrieved from <https://transcripts.factcheck.org/state-of-the-union-address/>
36. Trump Letter to Pelosi on Impeachment • Transcripts . FactCheck.org. (2019, December 17). Retrieved from <https://transcripts.factcheck.org/trump-letter-to-pelosi-on-impeachment/>
37. Villa, L. (2019, October 26). Can Bernie Sanders Grow His Base? Retrieved from <https://time.com/5711364/bernie-sanders-iowa-grow-base/>

Appendix II – R Code

```

library(tm)
library(scales)
library(plotly)
library(igraph)
library(ggraph)
library(tidytext)
library(pdftools)
library(reshape2)
library(tidyverse)
library(textreadr)
library(wordcloud)

setwd("/Users/jaisankerv/Downloads/02 Individual Assignment/Data")
x <- list.files(path="/Users/jaisankerv/Downloads/02 Individual Assignment/Data", pattern =
"pdf$")

opi <- lapply(x, pdf_text)

lapply(opi, length)
corps <- Corpus(URISource(x),readerControl = list(reader = readPDF))
my_df <- tidy(corps)

junk_common<-
data_frame(word=c("donald","trump","president","applause","people","country","american",
"americans","america",
"united","god","bless",
"laughter","lot","ahead","called","talking","dollars","day","time",
"happen","ago","love","tremendous","citizens","billion","usa",
"sanders","senator","sen","margaret","bernie","campaign",
"chris","cuomo","question","brennan","hayes"),
lexicon="junk")
my_df_tokens <- my_df %>%
  group_by(id)%>%
  unnest_tokens(word, text) %>%
  anti_join(stop_words)%>%
  anti_join(junk_common)%>%
  anti_join(stop_words) %>%
  filter(
    !str_detect(word, pattern = "[[:digit:]]"), # removes any words with numeric digits
    !str_detect(word, pattern = "[[:punct:]]"), # removes any remaining punctuations
    !str_detect(word, pattern = "\\b(\\.\\b)" # removes any remaining single letter words
  ) %>%
  count(word,sort=TRUE)

trump_tokens<-my_df_tokens%>%
  filter(id=="Trump.pdf")
bernie_tokens<-my_df_tokens%>%
  filter(id=="Bernie.pdf")

```

```

#-----FREQUENCY-----
ggplot(trump_tokens %>% filter(n >= 50), aes(x=reorder(word,n), y=n))+
  geom_col() + coord_flip() + ggtitle("Donald Trump")+
  geom_bar(stat = "identity", alpha = 0.7, fill = "#E91D0E") + #republican red #E91D0E
  theme_minimal(base_size = 14) +
  theme(panel.grid.major.y = element_blank(),
        panel.grid.minor.y = element_blank()) +
  ylab("frequency") +
  xlab("words")

ggplot(bernie_tokens %>% filter(n >= 40), aes(x=reorder(word,n), y=n))+
  geom_col() + coord_flip() + ggtitle("Bernie Sanders")+
  geom_bar(stat = "identity", alpha = 0.7, fill = "#0015BC") + #democratic blue #0015BC
  theme_minimal(base_size = 14) +
  theme(panel.grid.major.y = element_blank(),
        panel.grid.minor.y = element_blank()) +
  ylab("frequency") +
  xlab("words")

#-----AFINN SENTIMENT-----
trump_tokens %>% inner_join(get_sentiments("afinn")) %>% mutate((total= value * n))
%>% summarise(sum=sum(value))
bernie_tokens %>% inner_join(get_sentiments("afinn")) %>% mutate((total= value * n))
%>% summarise(sum=sum(value))

#-----BING SENTIMENT-----
trump_sentiments<-trump_tokens %>%
  inner_join(get_sentiments("bing"))%>%
  count(word,sentiment)%>%
  count(sentiment)

bernie_sentiments<-bernie_tokens %>%
  inner_join(get_sentiments("bing"))%>%
  count(word,sentiment)%>%
  count(sentiment)

trump_sentiments
bernie_sentiments

#-----SENTIMENT WORD CLOUD-----
-----

my_df_tokens_sentiments <- my_df %>%
  unnest_tokens(word, text) %>%
  anti_join(stop_words)%>%
  anti_join(junk_common)%>%
  inner_join(get_sentiments("nrc")) %>%
  count(word,sentiment)%>%
  count(sentiment)
my_df_tokens_sentiments

```

```

my_df_tokens_sentiments %>%
  inner_join(get_sentiments("nrc")) %>%
  count(word, sentiment, sort=TRUE) %>%
  acast(word ~sentiment, value.var="n", fill=0) %>%
  comparison.cloud(colors = c("grey20", "gray80"),
                    max.words=100, fixed.asp=TRUE, scale=c(0.8,0.8), title.size=1, rot.per=0.25)
#-----BI-GRAMS-----
bi_grams <- my_df %>%
  unnest_tokens(bigram, text, token = "ngrams", n = 2) %>%
  separate(bigram, into = c("first","second"), sep = " ", remove = FALSE) %>% # remove
stop words from tidytext package
  anti_join(stop_words, by = c("first" = "word")) %>%
  anti_join(junk_common, by = c("first" = "word")) %>%
  anti_join(stop_words, by = c("second" = "word")) %>%
  anti_join(junk_common, by = c("second" = "word")) %>%
  filter(str_detect(first, "[a-z]"),
         str_detect(second, "[a-z]")) %>%
  group_by(id) %>%
  count(bigram) %>%
  arrange(-n)
bi_grams_trump <- bi_grams %>%
  filter(id=="Trump.pdf")
bi_grams_bernies <- bi_grams %>%
  filter(id=="Bernie.pdf")

bi_grams_trump
bi_grams_bernies

#Cleaning
top_bigram_f$id <- gsub("Trump.pdf","Donald Trump", top_bigram_f$id)
top_bigram_f$id <- gsub("Bernie.pdf","Bernie Sanders", top_bigram_f$id)

bigram_f$id <- gsub("Trump.pdf","Donald Trump", bigram_f$id)
bigram_f$id <- gsub("Bernie.pdf","Bernie Sanders", bigram_f$id)

bigram_graph <- bigram_f %>%
  filter(n>7) %>%
  graph_from_data_frame()
bigram_graph
ggraph(bigram_graph, layout = "fr") +
  geom_edge_link()+
  geom_node_point()+
  geom_node_text(aes(label=name), vjust=1, hjust=1)

#-----tf-idf for bigram-----
#couldn't derive any meaning ful insights
bigram_tf_idf <- bi_grams %>%
  count(id, bigram) %>%
  bind_tf_idf(bigram, id, n) %>%

```

```

  arrange(desc(tf_idf))

bigram_tf_idf
bigram_tf_idf_trump<-bigram_tf_idf%>%
  filter(id=="Trump.pdf")

bigram_tf_idf_bernies<-bigram_tf_idf%>%
  filter(id=="Bernie.pdf")

#-----tf-idf-----
original_df <- my_df %>%
  unnest_tokens(word, text) %>%
  anti_join(stop_words) %>%
  anti_join(junk_common) %>%
  count(id, word, sort=TRUE) %>%
  ungroup()
total_words <- original_df %>%
  group_by(id) %>%
  summarize(total=sum(n))
original_df_words <- left_join(original_df, total_words)
print(original_df_words)
original_df_words <- original_df_words %>%
  bind_tf_idf(word, id, n)
original_df_words
original_df_words %>%
  arrange(desc(tf_idf))

original_df_words$id <- gsub("Trump.pdf","Donald Trump", original_df_words$id)
original_df_words$id <- gsub("Bernie.pdf","Bernie Sanders", original_df_words$id)

original_df_words %>%
  arrange(desc(tf_idf)) %>%
  mutate(word=factor(word, levels=rev(unique(word)))) %>%
  group_by(id) %>%
  top_n(30) %>%
  ungroup %>%
  ggplot(aes(word, tf_idf, fill=id))+
  geom_col(show.legend=FALSE)+
  labs(x=NULL, y="tf-idf")+
  theme_minimal(base_size = 14) +
  theme(panel.grid.major.y = element_blank(),
        panel.grid.minor.y = element_blank()) +
  facet_wrap(~id, ncol=2, scales="free")+
  coord_flip()

```

Appendix III – R Code Output

```

> lapply(opi, length)
[[1]]
[1] 100

[[2]]
[1] 217
a)

# A tibble: 1 x 2
  id          sum
  <chr>      <dbl>
1 Trump.pdf -167
> bernie_tokens %>% inn
Joining, by = "word"
# A tibble: 1 x 2
  id          sum
  <chr>      <dbl>
1 Bernie.pdf -190
b)

  id          sentiment      n
  <chr>      <chr>      <int>
1 Trump.pdf negative      555
2 Trump.pdf positive      400
> bernie_sentiments
# A tibble: 2 x 3
# Groups:   id [1]
  id          sentiment      n
  <chr>      <chr>      <int>
1 Bernie.pdf negative      371
2 Bernie.pdf positive      218
c)

```

```

      sentiment      n
      <chr>      <int>
1 anger          355
2 anticipation    274
3 disgust         219
4 fear           390
5 joy            216
5 negative        683
7 positive        691
3 sadness         315
3 surprise        148
3 trust          424
d) |

```

```

> bi_grams_trump
# A tibble: 5,861 x 3
# Groups:   id [1]
   id      bigram      n
   <chr>   <chr>   <int>
1 Trump.pdf supreme court    28
2 Trump.pdf tax reform       25
3 Trump.pdf fake news        23
4 Trump.pdf york times       22
5 Trump.pdf hillary clinton   20
6 Trump.pdf middle east      19
7 Trump.pdf law enforcement   18
8 Trump.pdf white house       17
9 Trump.pdf middle class      16
10 Trump.pdf north korea      16
# ... with 5,851 more rows
e) > |

```

```

> bi_grams_bernie
# A tibble: 3,082 x 3
# Groups:   id [1]
   id      bigram      n
   <chr>   <chr>   <int>
1 Bernie.pdf health care    53
2 Bernie.pdf wall street    30
3 Bernie.pdf climate change  23
4 Bernie.pdf minimum wage    20
5 Bernie.pdf hillary clinton  19
6 Bernie.pdf white house     18
7 Bernie.pdf criminal justice 17
8 Bernie.pdf fossil fuel     15
9 Bernie.pdf democratic party 14
10 Bernie.pdf tax breaks     14
# ... with 3,072 more rows
f) > |

```

```

> bigram_graph
IGRAPH 3a524b8 DN-- 78 80 --
+ attr: name (v/c), n (e/n), total (e/n), percent (e/n)
+ edges from 3a524b8 (vertex names):
[1] Bernie Sanders->health care      Bernie Sanders->wall street      Donald Trump ->supreme court
[4] Donald Trump ->tax reform          Bernie Sanders->climate change    Donald Trump ->fake news
[7] Donald Trump ->york times           Bernie Sanders->minimum wage      Donald Trump ->hillary clinton
[10] Bernie Sanders->hillary clinton    Donald Trump ->middle east        Bernie Sanders->white house
[13] Donald Trump ->law enforcement      Bernie Sanders->criminal justice   Donald Trump ->white house
[16] Donald Trump ->middle class         Donald Trump ->north korea        Donald Trump ->tax rate
[19] Donald Trump ->what's happening    Bernie Sanders->fossil fuel       Donald Trump ->southern border
[22] Donald Trump ->tax code            Bernie Sanders->democratic party   Bernie Sanders->tax breaks
+ ... omitted several edges

```

g)

```

> bigram_tf_idf
# A tibble: 8,943 x 6
# Groups:   id [2]
  id      bigram      n      tf  idf  tf_idf
  <chr>    <chr>    <int>  <dbl> <dbl> <dbl>
1 Bernie.pdf 1,856th pledged 1 0.000324 0.693 0.000225
2 Bernie.pdf abiding faith 1 0.000324 0.693 0.000225
3 Bernie.pdf abiding worry 1 0.000324 0.693 0.000225
4 Bernie.pdf abraham joshua 1 0.000324 0.693 0.000225
5 Bernie.pdf abroad paying 1 0.000324 0.693 0.000225
6 Bernie.pdf absentee ballot 1 0.000324 0.693 0.000225
7 Bernie.pdf absentee ballots 1 0.000324 0.693 0.000225
8 Bernie.pdf absolute outrage 1 0.000324 0.693 0.000225
9 Bernie.pdf absolutely absolutely 1 0.000324 0.693 0.000225
10 Bernie.pdf absolutely confident 1 0.000324 0.693 0.000225
# ... with 8,933 more rows

```

h)

```

> original_df_words %>%
+   arrange(desc(tf_idf))
# A tibble: 8,962 x 7
  id      word      n total      tf  idf  tf_idf
  <chr>    <chr>    <int> <int>  <dbl> <dbl> <dbl>
1 Trump.pdf i've      132 27379 0.00482 0.693 0.00334
2 Trump.pdf we'll      76 27379 0.00278 0.693 0.00192
3 Bernie.pdf medicare   34 13923 0.00244 0.693 0.00169
4 Bernie.pdf billionaires 28 13923 0.00201 0.693 0.00139
5 Trump.pdf fake       48 27379 0.00175 0.693 0.00122
6 Bernie.pdf vermont    23 13923 0.00165 0.693 0.00115
7 Trump.pdf won't      40 27379 0.00146 0.693 0.00101
8 Bernie.pdf harris     20 13923 0.00144 0.693 0.000996
9 Trump.pdf regulations 39 27379 0.00142 0.693 0.000987
10 Trump.pdf they've    39 27379 0.00142 0.693 0.000987
# ... with 8,952 more rows

```

i)