Convention Text Analysis: Word Count

Jennifer Lin

11/25/2019

Contents

Introduction

The goal of this section is to understand the methods in analyzing text through the `quanteda` package in R. Here, I will analyze the appeals to morality for speakers in the 2016 RNC and DNC using a software resembling the Lingistic Inquiry and Word Count (LIWC) in R.

I chose this method as a reflection of those used by Jesse Graham and colleagues in their paper. From reading their article, they count the percentage of words that appeal to each foundation based on a dictionary they developed. This process reflects a simple word count of the texts and it's output without removing common words or stemming them.

To begin, I gathered the transcripts of the speeches from the 2016 DNC and RNC. The sources and data related to each speech can be located here: https://github.com/lin-jennifer/2016NCtranscripts. For each of the speeches, a text file was created and named in the format `[convention].[speakerlastname].txt`. For spakers with the same last name, such as the Trumps or the Clintons, or common last names such as "Smith", first names were introduced after the last name for clarity. Each unique file identifier can be matched with more details for the speech, including the night the speaker spoke, and the source of the transcript.

DNC Speeches

Creating the Corpus

To begin the analysis, I load in the data to create the corpus. To do this, ensure that the metadata file is in the same fiele path as all of the other speeches. This will make retreiving files simpler.

The instructions for how to do this are in the `quanteda` GitHub page. Here, I followed the instructions to create a corpus based on the State of the Union speech examples: https://github.com/quanteda/quanteda

```
# Load Packages
library("quanteda")
library(quanteda.corpora)
library("readtext")
library(quanteda.dictionaries)

# Create Corpus when reading in text
DNCCorpus <- corpus(readtext("~/Desktop/Working/Moral-Psychology/Conventions/DNC/dnc*.tx
    metacorpus = list(source = "http://www.presidency.ucsb.edu/sou.php"),
    docid_field = "doc_id")
```

Generating Output

In the text analysis section, I use the Moral Foundations Dictionary that comes with the `quanteda.dictionaries` package. To get a copy of the dictionary, simply clone the package GitHub: https://github.com/kbenoit/quanteda.dictionaries

```
DNC <- liwcalike(DNCCorpus, dictionary = data_dictionary_MFD)

head(DNC)
```

```
##             docname Segment      WPS   WC Sixltr  Dic care.virtue
## 1    dnc.akbari.txt       1 22.94444  413  17.92 8.23        1.21
## 2  dnc.albright.txt       2 21.67500  867  18.57 4.96        0.69
## 3 dnc.alexander.txt       3 13.10526  249  16.06 5.22        2.41
## 4 dnc.allenjohn.txt       4 19.55263  743  17.50 6.59        0.54
## 5     dnc.amaru.txt       5 19.59091  431  16.24 4.41        1.16
## 6     dnc.asian.txt       6 17.33333 1248  22.84 3.69        0.64
##   care.vice fairness.virtue fairness.vice loyalty.virtue loyalty.vice
## 1      1.45            1.21          0.24           1.69         0.00
## 2      0.92            0.35          0.00           1.38         0.00
## 3      0.80            0.00          0.00           1.20         0.00
## 4      0.27            0.13          0.00           2.96         0.13
```

```
## 5      0.23            0.23          0.00           0.93         0.00
## 6      0.48            0.40          0.08           1.68         0.00
##   authority.virtue authority.vice sanctity.virtue sanctity.vice AllPunc
## 1             1.94           0.00            0.48          0.00   14.77
## 2             1.61           0.00            0.00          0.00   14.42
## 3             0.80           0.00            0.00          0.00   18.88
## 4             2.15           0.27            0.13          0.00   12.11
## 5             0.93           0.00            0.46          0.46   14.15
## 6             0.40           0.00            0.00          0.00   13.46
##   Period Comma Colon SemiC QMark Exclam Dash Quote Apostro Parenth OtherP
## 1   2.18  5.81  0.00  0.73  1.21   0.97 0.73  3.15    2.18       0  14.04
## 2   4.61  4.84  0.12  0.00  0.00   0.00 2.88  1.96    1.50       0  11.53
## 3   7.23  4.42  0.00  0.00  0.40   0.00 0.80  6.02    2.01       0  18.07
## 4   5.11  4.71  0.67  0.00  0.00   0.00 1.62  0.00    0.00       0  10.50
## 5   4.87  7.42  0.00  0.00  0.00   0.23 0.46  1.16    1.16       0  13.69
## 6   5.29  3.45  0.72  0.00  0.00   0.72 1.76  1.52    1.36       0  11.70
```

RNC Speeches

Repeat the process like the DNC used above

```
# Create Cropus
RNCCorpus <- corpus(readtext("~/Desktop/Working/Moral-Psychology/Conventions/RNC/rnc*.tx
    metacorpus = list(source = "http://www.presidency.ucsb.edu/sou.php"),
    docid_field = "doc_id")

# Analyse text
RNC <- liwcalike(RNCCorpus, dictionary = data_dictionary_MFD)

head(RNC)
```

```
##             docname Segment      WPS   WC Sixltr  Dic care.virtue
## 1  rnc.alvarado.txt       1 20.89744  815  21.72 6.63        0.74
## 2      rnc.baio.txt       2 10.54348  485  13.61 4.74        0.62
## 3   rnc.barrack.txt       3 20.78000 1039  15.98 2.89        1.06
## 4 rnc.beardsley.txt       4 19.41667  466  21.67 5.36        0.21
## 5 rnc.blackburn.txt       5 15.38095  646  17.80 6.35        0.62
## 6     rnc.bondi.txt       6 15.18519  410  22.44 8.54        1.46
##   care.vice fairness.virtue fairness.vice loyalty.virtue loyalty.vice
## 1      0.12            0.49          0.25           1.72         0.12
## 2      0.21            0.21          0.00           2.06         0.00
## 3      0.29            0.19          0.00           0.38         0.00
## 4      0.64            0.21          0.00           1.29         0.64
## 5      0.15            0.15          0.00           2.32         0.00
```

```
## 6      0.73            3.66          0.00           0.49         0.24
##   authority.virtue authority.vice sanctity.virtue sanctity.vice AllPunc
## 1             1.47           0.12            1.47          0.12   13.62
## 2             1.03           0.00            0.62          0.00   16.70
## 3             0.67           0.10            0.19          0.00   13.76
## 4             1.93           0.00            0.43          0.00   11.59
## 5             2.79           0.00            0.31          0.00   17.49
## 6             1.22           0.49            0.24          0.00   14.88
##   Period Comma Colon SemiC QMark Exclam Dash Quote Apostro Parenth OtherP
## 1   4.66  7.36  0.49  0.12  0.12   0.00 0.37  0.25    0.25       0  13.01
## 2   9.48  2.27  0.41  0.00  0.21   0.00 1.24  3.09    2.68       0  15.46
## 3   4.72  5.68  0.19  0.00  0.10   0.10 0.58  2.21    1.06       0  13.19
## 4   4.94  4.08  0.00  0.00  0.21   0.00 0.64  1.50    1.07       0  10.94
## 5   5.88  7.28  0.15  0.00  0.31   0.31 0.62  2.94    2.63       0  16.87
## 6   6.34  3.66  0.73  0.00  0.24   0.00 0.00  2.93    2.93       0  14.88
```

Analysis of Results

For each of the outputs for the RNC and DNC word count, I expored the data into separate CSV files in my local script and combined the outputs for a composite set. I also merged it with the metadata file in the repository so that it can contain information related to each of the speeches.

The code for this process looked as follows (Here I use the DNC as an example)

```
#Load metadata
metadata <- read.csv("~/Desktop/Working/Moral-Psychology/Conventions/DNC
  /metadata.csv", stringsAsFactors = FALSE, header = TRUE)
metadata$docname <- metadata$Filename

#Merge DNC output with the

DNCexp <- merge(DNC2, metadata, by = "docname")

#Export
library(rio)
export(DNCexp, "DNC.csv")
```

Then, once I have the information for the RNC and DNC, I combined them into one joined CSV, which I will use for the analysis.

For the combined set, it was created as follows:

```
#Merge the data frames
composite <- rbind(DNC, RNC)
```

For the analysis, I will begin by loading in this completed composite data set

```
# Load data
speech <- read.csv("~/Desktop/Working/Moral-Psychology/SpeechAnalysis/quanteda/Composite
    header = TRUE)
```

Since the printout has a "virtue" and "vice" category, I will combine the scores for the categories to get one composite score for each moral foundation

```
speech$Harm <- rowSums(speech[, c("care.virtue", "care.vice")],
    na.rm = TRUE)
speech$Fairness <- rowSums(speech[, c("fairness.virtue", "fairness.vice")],
    na.rm = TRUE)
speech$Ingroup <- rowSums(speech[, c("loyalty.virtue", "loyalty.vice")],
    na.rm = TRUE)
speech$Authority <- rowSums(speech[, c("authority.virtue", "authority.vice")],
    na.rm = TRUE)
speech$Purity <- rowSums(speech[, c("sanctity.virtue", "sanctity.vice")],
    na.rm = TRUE)
```

In the analysis, I will sort the output by convention and compare them. I will describe the scores of the harm foundation in detain, and repeat the same process for each of the subsequent analyses.

Before I begin, I first must load some packages

```
library(psych)
library(effsize)
```

To understand the difference between the Republicans and Democrats in the Harm foundation, I run descriptive statistics for this foundation, sorted by the convention

```
describeBy(speech$Harm, speech$Convention)
```

```
##
##  Descriptive statistics by group
## group: DNC
##    vars   n mean   sd median trimmed mad min  max range skew kurtosis   se
## X1    1 152  1.6 1.09   1.43    1.49 0.9   0 7.39  7.39 1.54     4.65 0.09
## ------------------------------------------------------------
## group: RNC
##    vars  n mean   sd median trimmed  mad min  max range skew kurtosis   se
## X1    1 64  1.1 0.63      1    1.06 0.45   0 2.93  2.93 0.73     0.21 0.08
```

Next, I run t-tests and acquire a measure of effect size (Cohen's d) for the Harm foundation

```
t.test(speech$Harm ~ speech$Convention)
```

```
##
##  Welch Two Sample t-test
```

```
##
## data:  speech$Harm by speech$Convention
## t = 4.219, df = 193.26, p-value = 3.77e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.2669420 0.7356238
## sample estimates:
## mean in group DNC mean in group RNC
##          1.603158          1.101875
```

```
cohen.d(speech$Harm, speech$Convention)
```

```
##
## Cohen's d
##
## d estimate: 0.511537 (medium)
## 95 percent confidence interval:
##     lower     upper
## 0.2138426 0.8092313
```

This process is repeated for Fairness, Ingroup, Authority, and Purity

```
### Fairness ###
describeBy(speech$Fairness, speech$Convention)
```

```
##
##  Descriptive statistics by group
## group: DNC
##    vars   n mean   sd median trimmed  mad min max range skew kurtosis   se
## X1    1 152 0.53 0.52   0.36    0.45 0.53   0 2.4   2.4 1.14     0.91 0.04
## ------------------------------------------------------------
## group: RNC
##    vars  n mean   sd median trimmed  mad min  max range skew kurtosis   se
## X1    1 64 0.44 0.55   0.26    0.35 0.33   0 3.66  3.66 3.35    16.14 0.07
```

```
t.test(speech$Fairness ~ speech$Convention)
```

```
##
##  Welch Two Sample t-test
##
## data:  speech$Fairness by speech$Convention
## t = 1.0498, df = 113.84, p-value = 0.296
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.07507386  0.24435017
## sample estimates:
## mean in group DNC mean in group RNC
```

```
##         0.5265132         0.4418750
```

```
cohen.d(speech$Fairness, speech$Convention)
```

```
##
## Cohen's d
##
## d estimate: 0.1593042 (negligible)
## 95 percent confidence interval:
##      lower      upper
## -0.1347992  0.4534075
```

```
### Ingroup ###
describeBy(speech$Ingroup, speech$Convention)
```

```
##
##  Descriptive statistics by group
## group: DNC
##    vars   n mean   sd median trimmed  mad min  max range skew kurtosis
## X1    1 152 1.45 0.89   1.29    1.36 0.67   0 6.49  6.49 1.68      5.8
##      se
## X1 0.07
## ------------------------------------------------------------
## group: RNC
##    vars  n mean   sd median trimmed  mad min  max range skew kurtosis   se
## X1    1 64 1.32 0.74   1.28    1.28 0.73   0 3.89  3.89 0.75     1.16 0.09
```

```
t.test(speech$Ingroup ~ speech$Convention)
```

```
##
##  Welch Two Sample t-test
##
## data:  speech$Ingroup by speech$Convention
## t = 1.0972, df = 141.32, p-value = 0.2744
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1032657  0.3608643
## sample estimates:
## mean in group DNC mean in group RNC
##          1.448487          1.319687
```

```
cohen.d(speech$Ingroup, speech$Convention)
```

```
##
## Cohen's d
##
## d estimate: 0.1516765 (negligible)
## 95 percent confidence interval:
```

```
##      lower      upper
## -0.1423906  0.4457436
```

```
### Authority ###
describeBy(speech$Authority, speech$Convention)
```

```
##
##  Descriptive statistics by group
## group: DNC
##    vars   n mean   sd median trimmed  mad min  max range skew kurtosis
## X1    1 152 1.02 0.71    0.9    0.95 0.65   0 3.53  3.53  0.9      0.7
##      se
## X1 0.06
## ------------------------------------------------------------
## group: RNC
##    vars  n mean   sd median trimmed  mad min  max range skew kurtosis   se
## X1    1 64 1.33 0.72   1.24    1.28 0.68   0 3.28  3.28 0.66     0.08 0.09
```

```
t.test(speech$Authority ~ speech$Convention)
```

```
##
##  Welch Two Sample t-test
##
## data:  speech$Authority by speech$Convention
## t = -2.9658, df = 115.9, p-value = 0.003666
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.5291905 -0.1053983
## sample estimates:
## mean in group DNC mean in group RNC
##          1.017237          1.334531
```

```
cohen.d(speech$Authority, speech$Convention)
```

```
##
## Cohen's d
##
## d estimate: -0.4463398 (small)
## 95 percent confidence interval:
##      lower      upper
## -0.7430893 -0.1495903
```

```
### Purity ###
describeBy(speech$Purity, speech$Convention)
```

```
##
##  Descriptive statistics by group
## group: DNC
```

```
##    vars   n mean   sd median trimmed  mad min  max range skew kurtosis
## X1    1 152 0.54 0.87   0.24    0.35 0.36   0 6.48  6.48 3.43    15.82
##      se
## X1 0.07
## ------------------------------------------------------------
## group: RNC
##    vars  n mean   sd median trimmed  mad min  max range skew kurtosis   se
## X1    1 64 0.56 0.57   0.42    0.47 0.38   0 2.56  2.56 1.55     2.11 0.07
```

```
t.test(speech$Purity ~ speech$Convention)
```

```
##
##  Welch Two Sample t-test
##
## data:  speech$Purity by speech$Convention
## t = -0.27658, df = 177.96, p-value = 0.7824
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.2251148  0.1697694
## sample estimates:
## mean in group DNC mean in group RNC
##         0.5371711         0.5648437
```

```
cohen.d(speech$Purity, speech$Convention)
```

```
##
## Cohen's d
##
## d estimate: -0.034821 (negligible)
## 95 percent confidence interval:
##      lower      upper
## -0.3285546  0.2589126
```

Additional Analyses

From these results, and from studying the speeches that were gathered from the DNC and RNC, I was interested in what would happen if we removed the benefiction from the dataset. Given these are more likely to be full of appeals to religion, hence their purpose, removing them might allow us to gain a better perception of the extent to which regular speakers and politicians appeal to each of the foundations.

So to begin, I remove the benedictions

```
table(speech$Type)
```

```
##
```

```
## benediction      speech       video
##           7         208           1
```

```
speech <- speech[!(speech$Type == "benediction"), ]
```

This leaves us with 208 speeches and 1 video transcript (which I kept because it was from someone who could not attend in person and the words would have been similiar to what they would otherwise say on the real stage.

In the next section, I repeat all the analyses for each foundation given the new conditions

```
### Harm ###
describeBy(speech$Harm, speech$Convention)
```

```
##
##  Descriptive statistics by group
## group: DNC
##    vars   n mean   sd median trimmed  mad min  max range skew kurtosis
## X1    1 145 1.54 1.05   1.39    1.44 0.86   0 7.39  7.39 1.68     5.95
##      se
## X1 0.09
## ------------------------------------------------------------
## group: RNC
##    vars  n mean   sd median trimmed  mad min  max range skew kurtosis   se
## X1    1 64  1.1 0.63      1    1.06 0.45   0 2.93  2.93 0.73     0.21 0.08
```

```
t.test(speech$Harm ~ speech$Convention)
```

```
##
##  Welch Two Sample t-test
##
## data:  speech$Harm by speech$Convention
## t = 3.7191, df = 187.63, p-value = 0.000264
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.2055389 0.6698835
## sample estimates:
## mean in group DNC mean in group RNC
##          1.539586          1.101875
```

```
cohen.d(speech$Harm, speech$Convention)
```

```
##
## Cohen's d
##
## d estimate: 0.4646489 (small)
## 95 percent confidence interval:
##     lower     upper
```

```
## 0.1654104 0.7638875
```

```
### Fairness ###
describeBy(speech$Fairness, speech$Convention)
```

```
##
##  Descriptive statistics by group
## group: DNC
##    vars   n mean   sd median trimmed  mad min max range skew kurtosis   se
## X1    1 145 0.52 0.53   0.36    0.45 0.53   0 2.4   2.4 1.16     0.96 0.04
## ------------------------------------------------------------
## group: RNC
##    vars  n mean   sd median trimmed  mad min  max range skew kurtosis   se
## X1    1 64 0.44 0.55   0.26    0.35 0.33   0 3.66  3.66 3.35    16.14 0.07
```

```
t.test(speech$Fairness ~ speech$Convention)
```

```
##
##  Welch Two Sample t-test
##
## data:  speech$Fairness by speech$Convention
## t = 0.99746, df = 116.49, p-value = 0.3206
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.07992156  0.24210260
## sample estimates:
## mean in group DNC mean in group RNC
##         0.5229655         0.4418750
```

```
cohen.d(speech$Fairness, speech$Convention)
```

```
##
## Cohen's d
##
## d estimate: 0.151957 (negligible)
## 95 percent confidence interval:
##      lower      upper
## -0.1442708  0.4481848
```

```
### Ingroup ###
describeBy(speech$Ingroup, speech$Convention)
```

```
##
##  Descriptive statistics by group
## group: DNC
##    vars   n mean   sd median trimmed  mad min  max range skew kurtosis
## X1    1 145 1.42 0.87   1.23    1.33 0.68   0 6.49  6.49 1.81     6.88
##      se
```

```
## X1 0.07
## ------------------------------------------------------------
## group: RNC
##    vars  n mean   sd median trimmed  mad min  max range skew kurtosis   se
## X1    1 64 1.32 0.74   1.28    1.28 0.73   0 3.89  3.89 0.75     1.16 0.09
```

```
t.test(speech$Ingroup ~ speech$Convention)
```

```
##
##  Welch Two Sample t-test
##
## data:  speech$Ingroup by speech$Convention
## t = 0.81248, df = 140.81, p-value = 0.4179
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1368534  0.3278232
## sample estimates:
## mean in group DNC mean in group RNC
##          1.415172          1.319687
```

```
cohen.d(speech$Ingroup, speech$Convention)
```

```
##
## Cohen's d
##
## d estimate: 0.1144192 (negligible)
## 95 percent confidence interval:
##      lower      upper
## -0.1816516  0.4104900
```

```
### Authority ###
describeBy(speech$Authority, speech$Convention)
```

```
##
##  Descriptive statistics by group
## group: DNC
##    vars   n mean  sd median trimmed  mad min  max range skew kurtosis   se
## X1    1 145 1.01 0.7   0.91    0.95 0.65   0 3.53  3.53 0.91     0.84 0.06
## ------------------------------------------------------------
## group: RNC
##    vars  n mean   sd median trimmed  mad min  max range skew kurtosis   se
## X1    1 64 1.33 0.72   1.24    1.28 0.68   0 3.28  3.28 0.66     0.08 0.09
```

```
t.test(speech$Authority ~ speech$Convention)
```

```
##
##  Welch Two Sample t-test
##
```

```
## data:  speech$Authority by speech$Convention
## t = -2.9837, df = 116.71, p-value = 0.003469
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.5328302 -0.1076806
## sample estimates:
## mean in group DNC mean in group RNC
##          1.014276          1.334531
```

```
cohen.d(speech$Authority, speech$Convention)
```

```
##
## Cohen's d
##
## d estimate: -0.4541739 (small)
## 95 percent confidence interval:
##      lower      upper
## -0.7532629 -0.1550849
```

```
### Purity ###
describeBy(speech$Purity, speech$Convention)
```

```
##
##  Descriptive statistics by group
## group: DNC
##    vars   n mean   sd median trimmed  mad min  max range skew kurtosis
## X1    1 145 0.41 0.55   0.23    0.31 0.34   0 3.12  3.12 2.26     6.61
##      se
## X1 0.05
## ------------------------------------------------------------
## group: RNC
##    vars  n mean   sd median trimmed  mad min  max range skew kurtosis   se
## X1    1 64 0.56 0.57   0.42    0.47 0.38   0 2.56  2.56 1.55     2.11 0.07
```

```
t.test(speech$Purity ~ speech$Convention)
```

```
##
##  Welch Two Sample t-test
##
## data:  speech$Purity by speech$Convention
## t = -1.8233, df = 117.69, p-value = 0.0708
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.32029268  0.01322587
## sample estimates:
## mean in group DNC mean in group RNC
##         0.4113103         0.5648437
```

```r
cohen.d(speech$Purity, speech$Convention)
```

```
##
## Cohen's d
##
## d estimate: -0.2765204 (small)
## 95 percent confidence interval:
##       lower       upper
## -0.57358468  0.02054388
```

Graph of Results

I generate a graph to display the results of the analyses. These do not contain the invocation and benediction addresses.

Before I begin, I load some packages

```r
# Load packages
library(car)
library(dplyr)
library(psych)
library(ggplot2)
library(GGally)
library("ggpubr")
library("reshape2")
library(scales)
```

I will display my data using bar graphs. To do this, I generate some summary statistics for each of the foundations

```r
# Summary Statistice by convention

Harm <- speech %>% group_by(Convention) %>% summarize(mean = mean(Harm,
    na.rm = TRUE), sd = sd(Harm, na.rm = TRUE), n = n(), se = sd/sqrt(n),
    ci = qt(0.975, df = n - 1) * se, max = max(Harm), min = min(Harm),
    med = median(Harm)) %>% mutate(type = "Harm")

Fairness <- speech %>% group_by(Convention) %>% summarize(mean = mean(Fairness,
    na.rm = TRUE), sd = sd(Fairness, na.rm = TRUE), n = n(),
    se = sd/sqrt(n), ci = qt(0.975, df = n - 1) * se, max = max(Fairness),
    min = min(Fairness), med = median(Fairness)) %>% mutate(type = "Fairness")

Ingroup <- speech %>% group_by(Convention) %>% summarize(mean = mean(Ingroup,
    na.rm = TRUE), sd = sd(Ingroup, na.rm = TRUE), n = n(), se = sd/sqrt(n),
    ci = qt(0.975, df = n - 1) * se, max = max(Ingroup), min = min(Ingroup),
```

```
    med = median(Ingroup)) %>% mutate(type = "Ingroup")

Authority <- speech %>% group_by(Convention) %>% summarize(mean = mean(Authority,
    na.rm = TRUE), sd = sd(Authority, na.rm = TRUE), n = n(),
    se = sd/sqrt(n), ci = qt(0.975, df = n - 1) * se, max = max(Authority),
    min = min(Authority), med = median(Authority)) %>% mutate(type = "Authority")

Purity <- speech %>% group_by(Convention) %>% summarize(mean = mean(Purity,
    na.rm = TRUE), sd = sd(Purity, na.rm = TRUE), n = n(), se = sd/sqrt(n),
    ci = qt(0.975, df = n - 1) * se, max = max(Purity), min = min(Purity),
    med = median(Purity)) %>% mutate(type = "Purity")
```

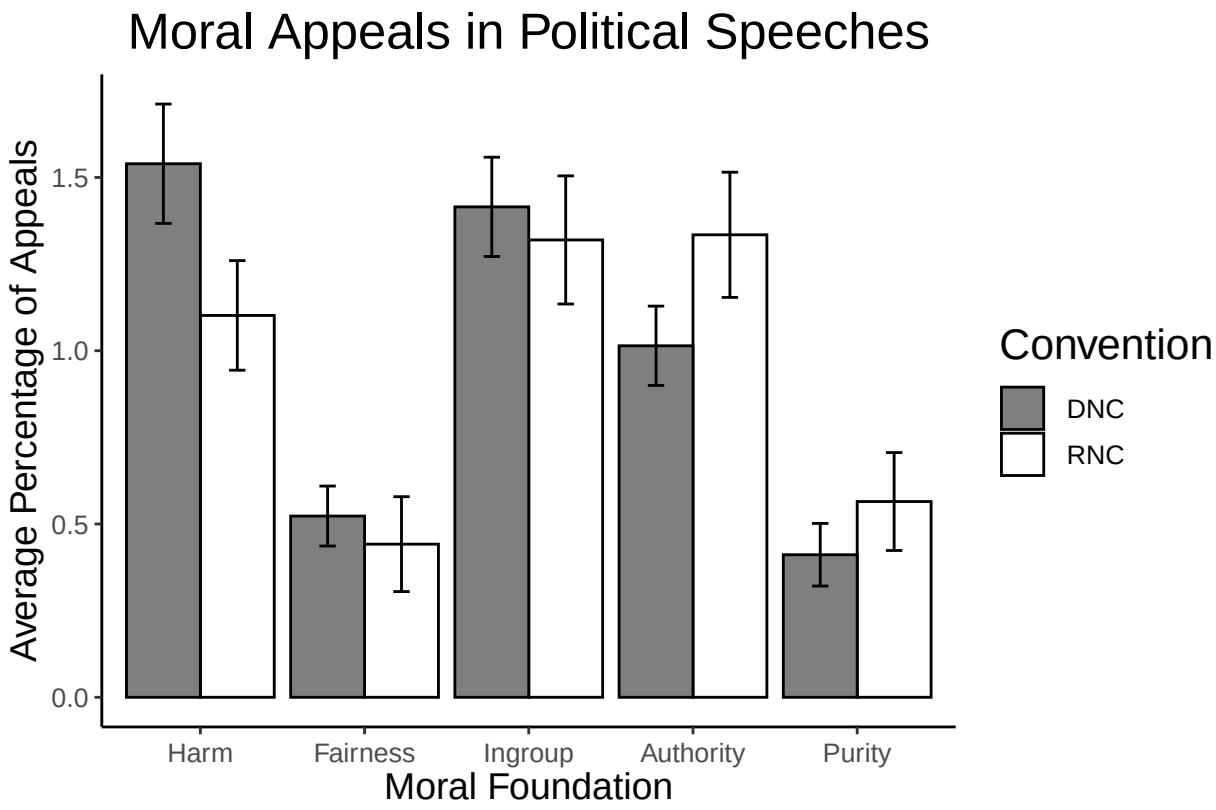
To generate the graphs, I need to bind these summary statistics to one dataframe

```
# Combine each of the outputs
token <- rbind(Harm, Fairness, Ingroup, Authority, Purity)

# Organize label order for the foundations
token$type <- factor(token$type, levels = c("Harm", "Fairness",
    "Ingroup", "Authority", "Purity"))
```

I generate the plot using ggplot

```
ggplot(token, aes(x = type, y = mean, fill = Convention)) + geom_bar(stat = "identity",
    position = position_dodge(), color = "black") + geom_errorbar(aes(ymin = mean -
    ci, ymax = mean + ci), width = 0.2, position = position_dodge(0.9)) +
    ggtitle("Moral Appeals in Political Speeches") + theme_classic() +
    xlab("Moral Foundation") + ylab("Average Percentage of Appeals") +
    labs(caption = "Source: 2016 RNC and DNC") + theme(text = element_text(size = 12,
    colour = "black"), axis.title = element_text(size = 14, colour = "black"),
    title = element_text(size = 16, colour = "black"), plot.caption = element_text(size
        color = "black"), axis.text.x = element_text(angle = 0,
        hjust = 0.5, vjust = 0.5), plot.title = element_text(hjust = 0.5)) +
    scale_x_discrete(labels = wrap_format(10)) + scale_fill_manual("Convention",
    values = c(DNC = "grey50", RNC = "white"))
```



Source: 2016 RNC and DNC