

510 CASES

Rachel Martin

9/11/2019

```
library(plyr)
library(mudata2)
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:plyr':
##
##   arrange, count, desc, failwith, id, mutate, rename, summarise,
##   summarize

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

library(ggplot2)
library(ggExtra)
library(ggribes)
library(tidyverse)

## — Attaching packages

tidyverse 1.2.1 —

## ✓ tibble 2.1.3      ✓ purrr 0.3.2
## ✓ tidyr 0.8.3       ✓ stringr 1.4.0
## ✓ readr 1.3.1      ✓ forcats 0.4.0

## — Conflicts

tidyverse_conflicts()

## ✗ dplyr::arrange() masks plyr::arrange()
## ✗ purrr::compact() masks plyr::compact()
## ✗ dplyr::count() masks plyr::count()
## ✗ dplyr::failwith() masks plyr::failwith()
## ✗ dplyr::filter() masks stats::filter()
## ✗ dplyr::id() masks plyr::id()
```

```
## ✗ dplyr::lag()          masks stats::lag()
## ✗ dplyr::mutate()       masks plyr::mutate()
## ✗ dplyr::rename()      masks plyr::rename()
## ✗ dplyr::summarise()    masks plyr::summarise()
## ✗ dplyr::summarize()    masks plyr::summarize()
```

```
library(na.tools)
```

```
cases <- read.csv("cases.csv")
```

HYPOTHESIS 1 MEASURES

Article mentions

```
all.sheet <- cases %>% select(RCsheet, MDsheet, JHsheet, DKsheet, BOSheet,
CPsheet, JRSheet, JSsheet, BSsheet, MTsheet, NWSheet)
```

```
colSums(all.sheet == "yes")
```

```
## RCsheet MDsheet JHsheet DKsheet BOSheet CPsheet JRSheet JSsheet BSsheet
##      187      284      104      138      193      184      158      278      145
## MTsheet NWSheet
##      299      121
```

```
names <- c("Corday", "DeWine", "Husted", "Kucinich", "O'Neill", "Pillich",
"Renacci", "Schiavoni", "Sutton", "Taylor", "Whaley")
```

```
names
```

```
## [1] "Corday"      "DeWine"      "Husted"      "Kucinich"    "O'Neill"
## [6] "Pillich"     "Renacci"     "Schiavoni"   "Sutton"      "Taylor"
## [11] "Whaley"
```

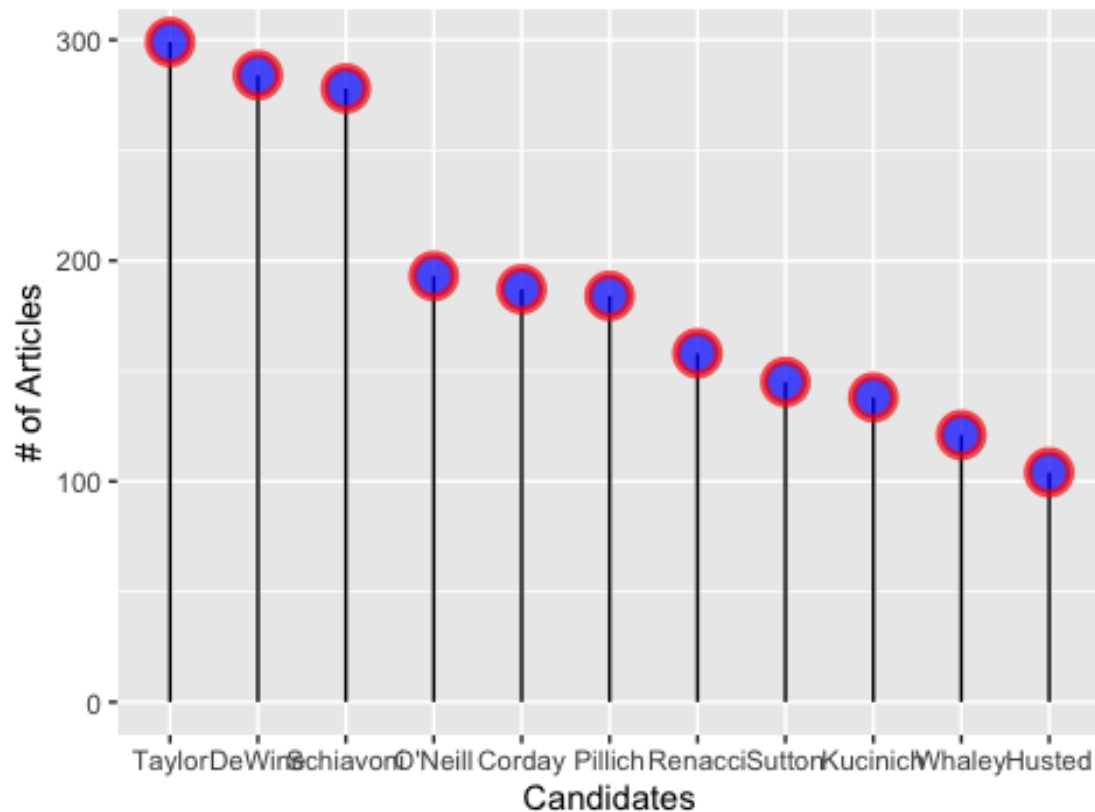
```
article <- c(187, 284, 104, 138, 193, 184, 158, 278, 145, 299, 121)
article
```

```
## [1] 187 284 104 138 193 184 158 278 145 299 121
```

```
df <- data.frame(names = names, article = article)
```

```
df <- data.frame(names = names, article = article)
ggplot(df, aes(x=reorder(names, -article), y=article)) +
  geom_segment(aes(x=reorder(names, -article), xend=names, y=0,
yend=article)) +
  geom_point(size=5, color="red", fill=alpha("blue", 0.3), alpha=0.7,
shape=21, stroke=2) + labs(x = "Candidates", y = "# of Articles", title =
"Article Quantity Among Candidates")
```

Article Quantity Among Candidates



Sentence quantity

```
all.sen <- cases %>% select(RCnum_sen, MDnum_sen, JHnum_sen, DKnum_sen,
  BOnum_sen, CPnum_sen, JRnum_sen, JSnum_sen, BSnum_sen, MTnum_sen, NWnum_sen)
all.sen[is.na(all.sen)] <- 0
```

```
colSums(all.sen)
```

```
## RCnum_sen MDnum_sen JHnum_sen DKnum_sen BOnum_sen CPnum_sen JRnum_sen
##      1490      2043       586      1059      1285       519       653
## JSnum_sen BSnum_sen MTnum_sen NWnum_sen
##      886      395      2006       465
```

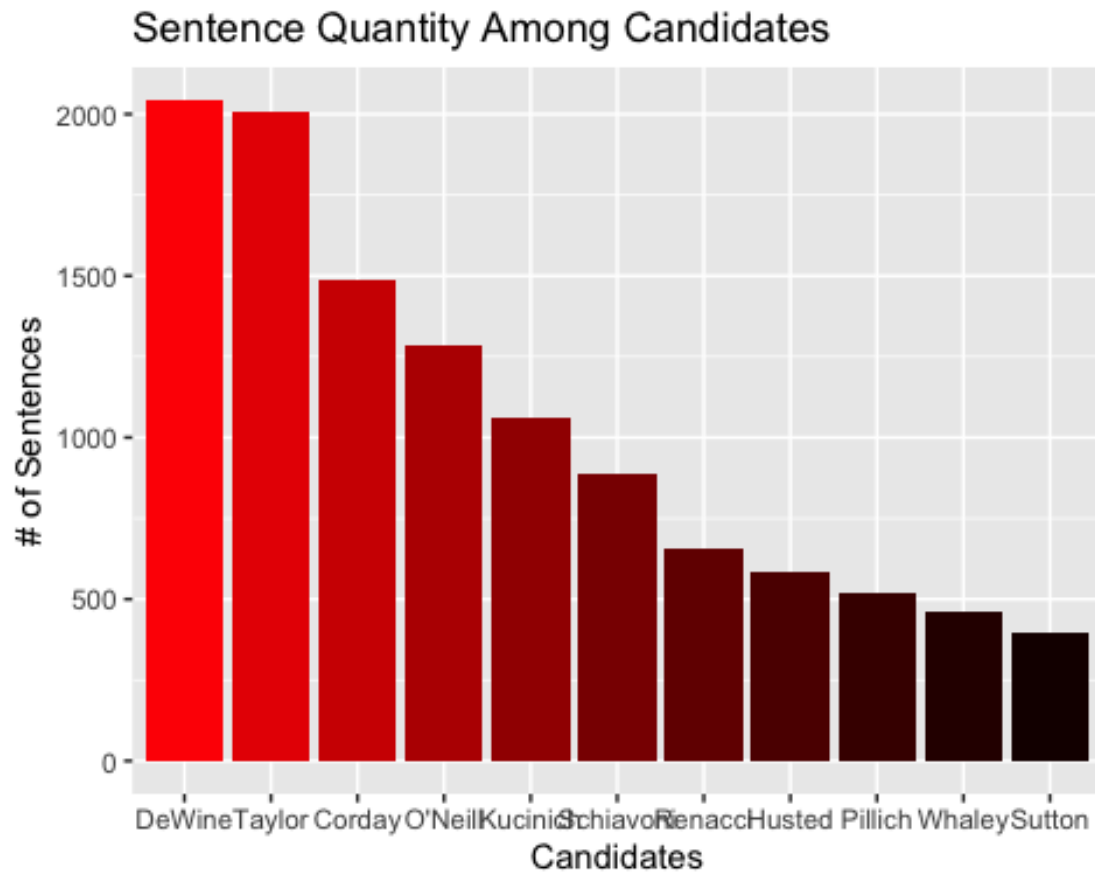
```
names
```

```
## [1] "Corday" "DeWine" "Husted" "Kucinich" "O'Neill"
## [6] "Pillich" "Renacci" "Schiavoni" "Sutton" "Taylor"
## [11] "Whaley"
```

```
sen <- c(1490, 2043, 586, 1059, 1285, 519, 653, 886, 395, 2006, 465)
sen
```

```
## [1] 1490 2043 586 1059 1285 519 653 886 395 2006 465
```

```
df1 <- data.frame(names = names, sen = sen)
ggplot(df1, aes(x=reorder(names, -sen), y=sen)) +
  geom_col(fill=rgb(11:1/11,0,0),) + labs(x = "Candidates", y = "# of
Sentences", title = "Sentence Quantity Among Candidates")
```



headline mentions

```
all.head <- cases %>% select(RChead, MDhead, JHhead, DKhead, BOhead, CPhead,
JRhead, JShead, BShead, MThead, NWhead)
```

```
table(all.head$RChead)
```

```
##
## no yes
## 118 69
```

```
table(all.head$MDhead)
```

```
##
## no yes
## 196 88
```

```
table(all.head$JHhead)
```

```
##
## no yes
## 81 23

table(all.head$DKhead)

##
## no yes
## 87 51

table(all.head$B0head)

##
## no yes
## 146 47

table(all.head$CPhead)

##
## no yes
## 162 22

table(all.head$JRhead)

##
## no yes
## 131 27

table(all.head$JShead)

##
## no yes
## 251 27

table(all.head$BShead)

##
## no yes
## 128 17

table(all.head$MThead)

##
## no yes
## 219 80

table(all.head$NWhead)

##
## no yes
## 105 17

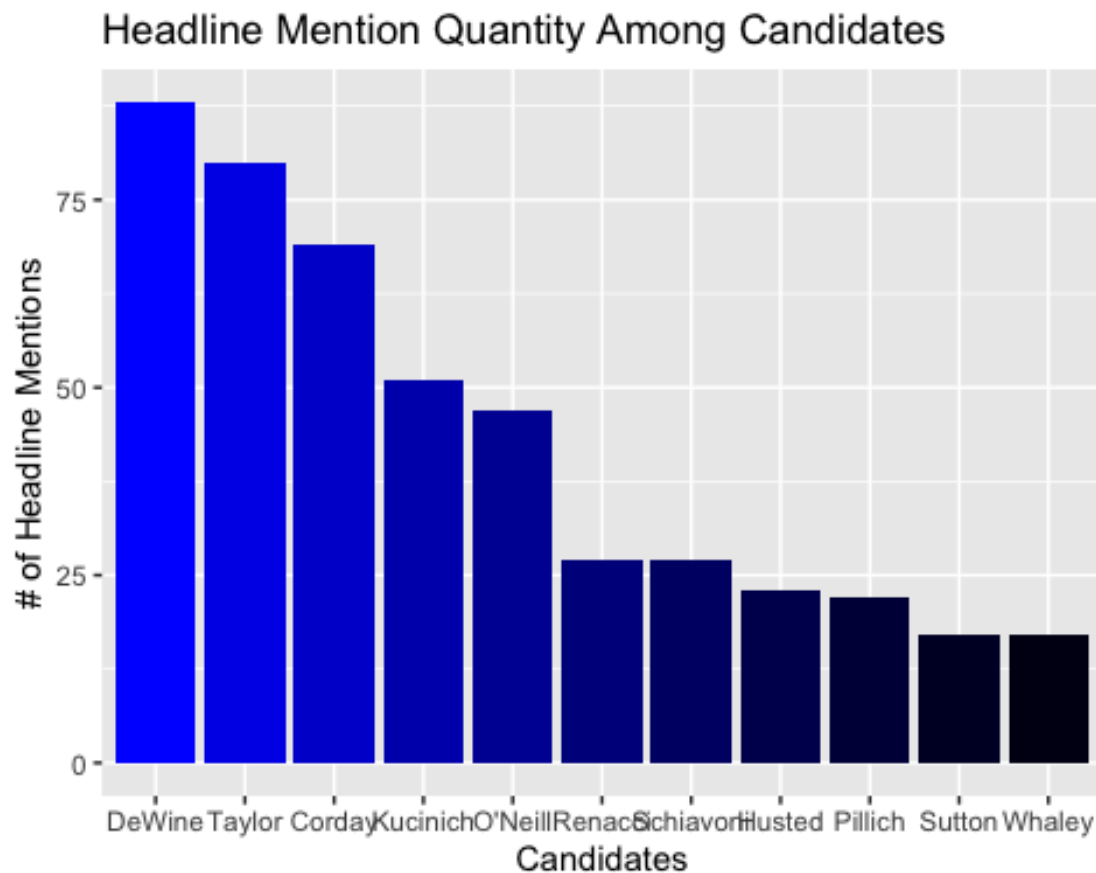
names
```

```
## [1] "Corday"      "DeWine"      "Husted"      "Kucinich"    "O'Neill"
## [6] "Pillich"     "Renacci"     "Schiavoni"   "Sutton"      "Taylor"
## [11] "Whaley"

head <- c(69, 88, 23, 51, 47, 22, 27, 27, 17, 80, 17)
head

## [1] 69 88 23 51 47 22 27 27 17 80 17

df4 <- data.frame(names = names, head = head)
ggplot(df4, aes(x=reorder(names, -head), y=head)) + geom_col(fill = rgb(0, 0, 11:1/11)) + labs(x = "Candidates", y = "# of Headline Mentions", title = "Headline Mention Quantity Among Candidates")
```



HYPOTHESIS 2 MEASURES

tone of articles

```
all.tone <- cases %>% select(RCtone, MDtone, JHtone, DKtone, BOTone, CPtone,
JRtone, JStone, BStone, MTtone, Nwtone)
table(all.tone$RCtone)
```

```

##
##  1  2  3
## 58 18 110

table(all.tone$MDtone)

##
##  1  2  3
## 64 37 183

table(all.tone$JHtone)

##
##  1  2  3
## 26  5 73

table(all.tone$DKtone)

##
##  1  2  3
## 25 20 93

table(all.tone$B0tone)

##
##  1  2  3
##  9 46 138

table(all.tone$CPtone)

##
##  1  2  3
## 17  4 163

table(all.tone$JRTone)

##
##  1  2  3
## 17  6 135

table(all.tone$JStone)

##
##  1  2  3
## 28  4 246

table(all.tone$BStone)

##
##  1  2  3
## 15  4 126

table(all.tone$MTtone)

```

```
##
##  1  2  3
## 44 31 224

table(all.tone$Nwtone)

##
##  1  2  3
## 16  1 105

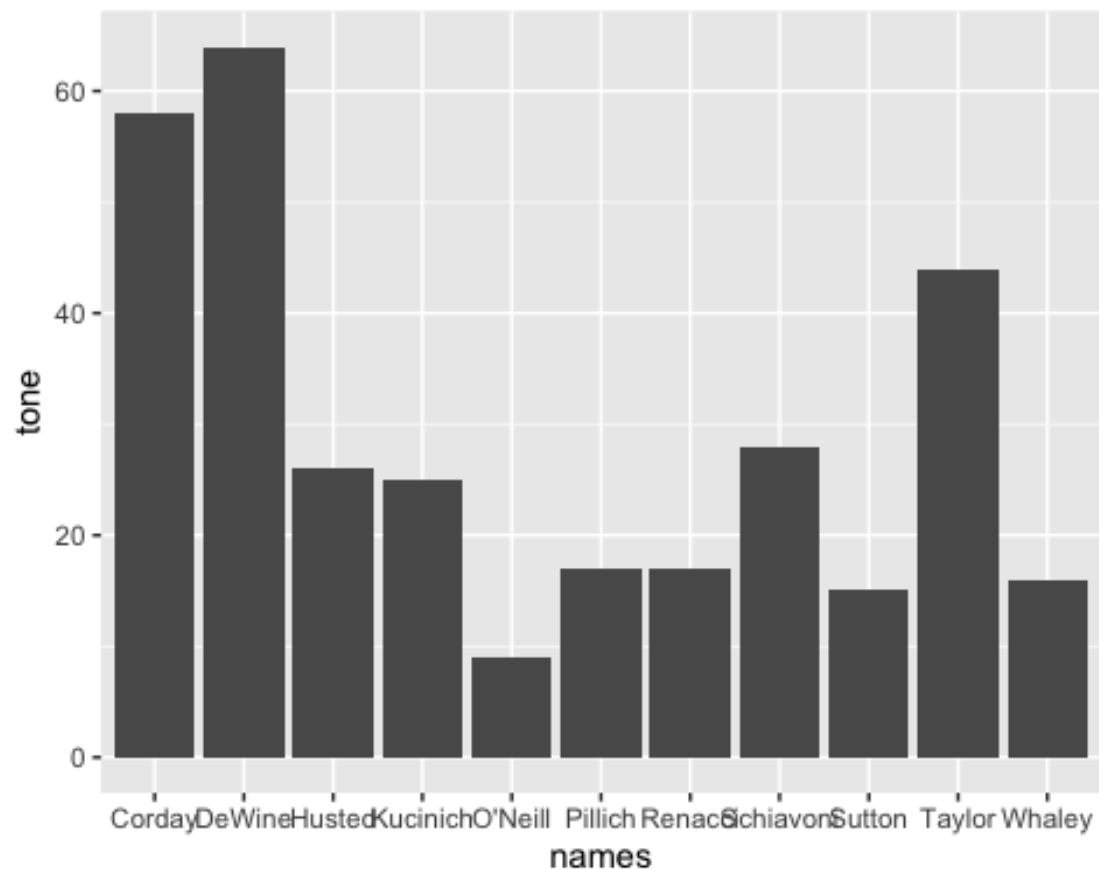
names

## [1] "Corday"    "DeWine"    "Husted"    "Kucinich"  "O'Neill"
## [6] "Pillich"   "Renacci"   "Schiavoni" "Sutton"    "Taylor"
## [11] "Whaley"

tone <- c(58, 64, 26, 25, 9, 17, 17, 28, 15, 44, 16)
tone

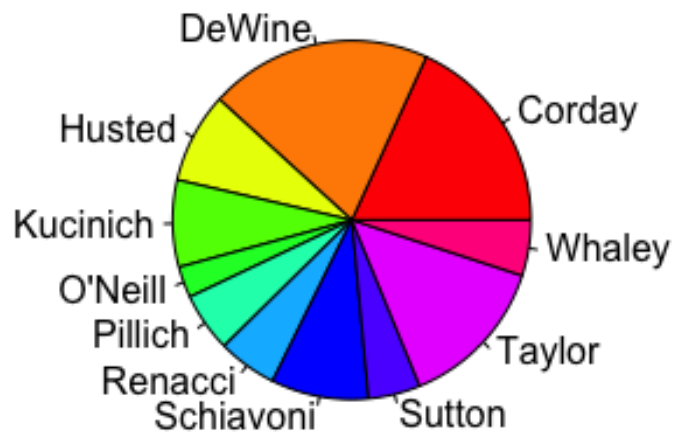
## [1] 58 64 26 25  9 17 17 28 15 44 16

df5 <- data.frame(names = names, tone = tone)
ggplot(df5, aes(names, tone)) + geom_col()
```




```
pie(tone, names, main = "Positive Tone Among Candidates", col =
rainbow(length(tone)))
```

Positive Tone Among Candidates



headline tone of articles (only out of articles mentioned and mentioned in the headline)

```
all.head.tone <- cases %>% select(RChead_tone, MDhead_tone, JHhead_tone,
DKhead_tone, BOhead_tone, CPhead_tone, JRhead_tone, JShead_tone, BShead_tone,
MThead_tone, NWhead_tone)
table(all.head.tone$RChead_tone)
```

```
##
##          Negative Neutral/Both Positive & Negative
##              11                      32
##          Positive
##              26
```

```
table(all.head.tone$MDhead_tone)
```

```
##
##          Negative Neutral/Both Positive & Negative
##              25                      39
##          Positive
##              24
```

```
table(all.head.tone$JHhead_tone)
```

```
##
##           Negative Neutral/Both Positive & Negative
##                5                                13
##           Positive
##                5
```

```
table(all.head.tone$DKhead_tone)
```

```
##
##           Negative Neutral/Both Positive & Negative
##                14                                25
##           Positive
##                12
```

```
table(all.head.tone$BOhead_tone)
```

```
##
##           Negative Neutral/Both Positive & Negative
##                21                                25
##           Positive
##                1
```

```
table(all.head.tone$CPhead_tone)
```

```
##
##           Negative Neutral/Both Positive & Negative
##                4                                14
##           Positive
##                4
```

```
table(all.head.tone$JRhead_tone)
```

```
##
##           Negative Neutral/Both Positive & Negative
##                4                                21
##           Positive
##                2
```

```
table(all.head.tone$JShead_tone)
```

```
##
##           Negative Neutral/Both Positive & Negative
##                3                                18
##           Positive
##                6
```

```
table(all.head.tone$BShead_tone)
```

```
##
##           Negative Neutral/Both Positive & Negative
##                1                                11
```

```
##                               Positive
##                               5

table(all.head.tone$MThead_tone)

##
##                               Negative Neutral/Both Positive & Negative
##                               30 35
##                               Positive
##                               15

table(all.head.tone$NWhead_tone)

##
##                               Negative Neutral/Both Positive & Negative
##                               1 11
##                               Positive
##                               5

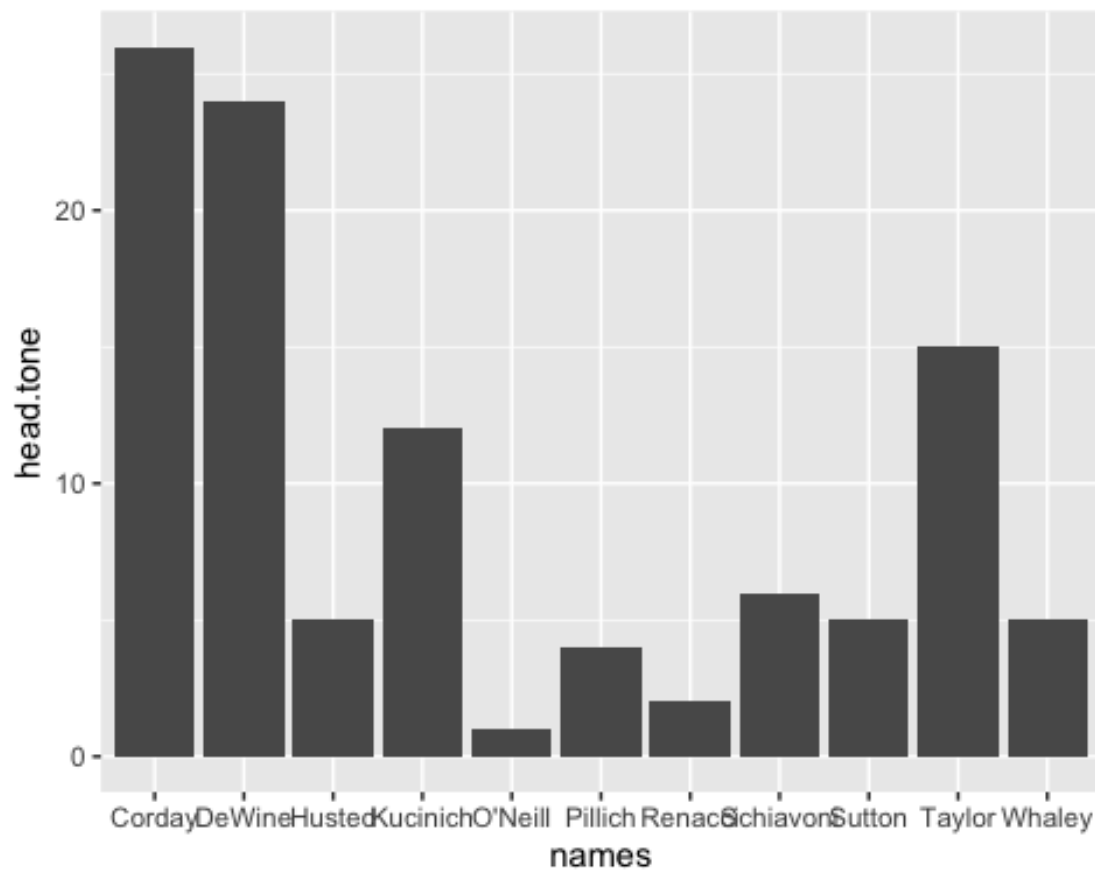
names

## [1] "Corday"    "DeWine"    "Husted"    "Kucinich"  "O'Neill"
## [6] "Pillich"    "Renacci"   "Schiavoni" "Sutton"    "Taylor"
## [11] "Whaley"

head.tone <- c(26, 24,5,12,1,4,2,6,5,15,5)
head.tone

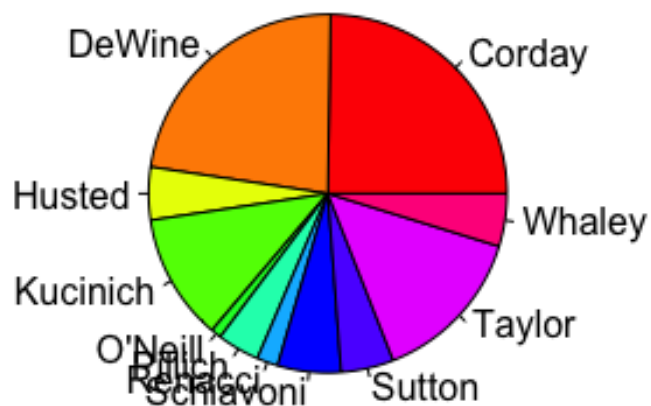
## [1] 26 24 5 12 1 4 2 6 5 15 5

df4 <- data.frame(names = names, head.tone = head.tone)
ggplot(df4, aes(names, head.tone)) + geom_col()
```



```
pie(head.tone, names, main = "Positive Headline Tone Among Candidates", col =  
rainbow(length(tone)))
```

Positive Headline Tone Among Candidates



HYPOTHESIS 3 MEASURES

qualifications

```
all.qual <- cases %>% select(RCqual, MDqual, JHqual, DKqual, BOqual, CPqual, JRqual, JSqual, BSqual, MTqual, NWqual)
table(all.qual$RCqual)
```

```
##
## no yes
## 77 110
```

```
table(all.qual$MDqual)
```

```
##
## no yes
## 193 91
```

```
table(all.qual$JHqual)
```

```
##
## no yes
## 78 26
```

```
table(all.qual$DKqual)
```

```
##
```

```
## no yes
```

```
## 69 69
```

```
table(all.qual$BQqual)
```

```
##
```

```
## no yes
```

```
## 137 56
```

```
table(all.qual$CPqual)
```

```
##
```

```
## no yes
```

```
## 161 23
```

```
table(all.qual$JRqual)
```

```
##
```

```
## no yes
```

```
## 126 32
```

```
table(all.qual$JSqual)
```

```
##
```

```
## no yes
```

```
## 245 33
```

```
table(all.qual$BSqual)
```

```
##
```

```
## no yes
```

```
## 126 19
```

```
table(all.qual$MTqual)
```

```
##
```

```
## no yes
```

```
## 226 73
```

```
table(all.qual$NWqual)
```

```
##
```

```
## no yes
```

```
## 109 13
```

```
names
```

```
## [1] "Corday" "DeWine" "Husted" "Kucinich" "O'Neill"
```

```
## [6] "Pillich" "Renacci" "Schiavoni" "Sutton" "Taylor"
```

```
## [11] "Whaley"
```

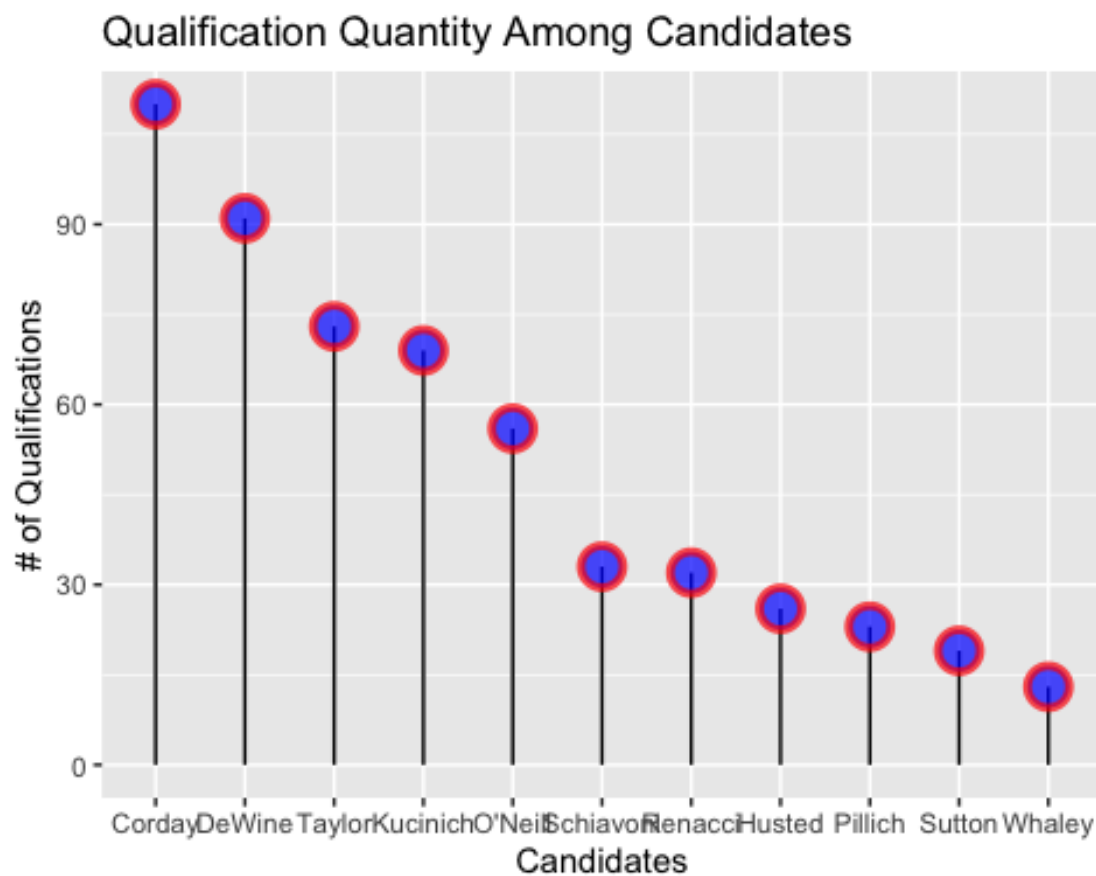
```

qual <- c(110, 91, 26, 69, 56, 23, 32, 33, 19, 73, 13)
qual

## [1] 110  91  26  69  56  23  32  33  19  73  13

df2 <- data.frame(names = names, qual = qual)
ggplot(df2, aes(x=reorder(names, -qual), y=qual)) +
  geom_segment(aes(x=reorder(names, -qual), xend=names, y=0, yend=qual)) +
  geom_point(size=5, color="red", fill=alpha("blue", 0.3), alpha=0.7,
shape=21, stroke=2) + labs(x = "Candidates", y = "# of Qualifications", title
= "Qualification Quantity Among Candidates")

```



HYPOTHESIS 4 MEASURES

positive traits

```

all.pos <- cases %>% select(RCpos, MDpos, JHpos, DKpos, B0pos, CPpos, JRpos,
JSpos, BSp0s, MTpos, NWpos)
table(all.pos$RCpos)

##
## no yes
## 81 106

table(all.pos$MDpos)

```

```
##
## no yes
## 144 140

table(all.pos$JHpos)

##
## no yes
## 69 35

table(all.pos$DKpos)

##
## no yes
## 91 47

table(all.pos$B0pos)

##
## no yes
## 171 22

table(all.pos$CPpos)

##
## no yes
## 155 29

table(all.pos$JRpos)

##
## no yes
## 121 37

table(all.pos$JSp0s)

##
## no yes
## 224 54

table(all.pos$BSpos)

##
## no yes
## 124 21

table(all.pos$MTpos)

##
## no yes
## 199 100

table(all.pos$NWpos)
```



```
##
## no yes
## 102 20

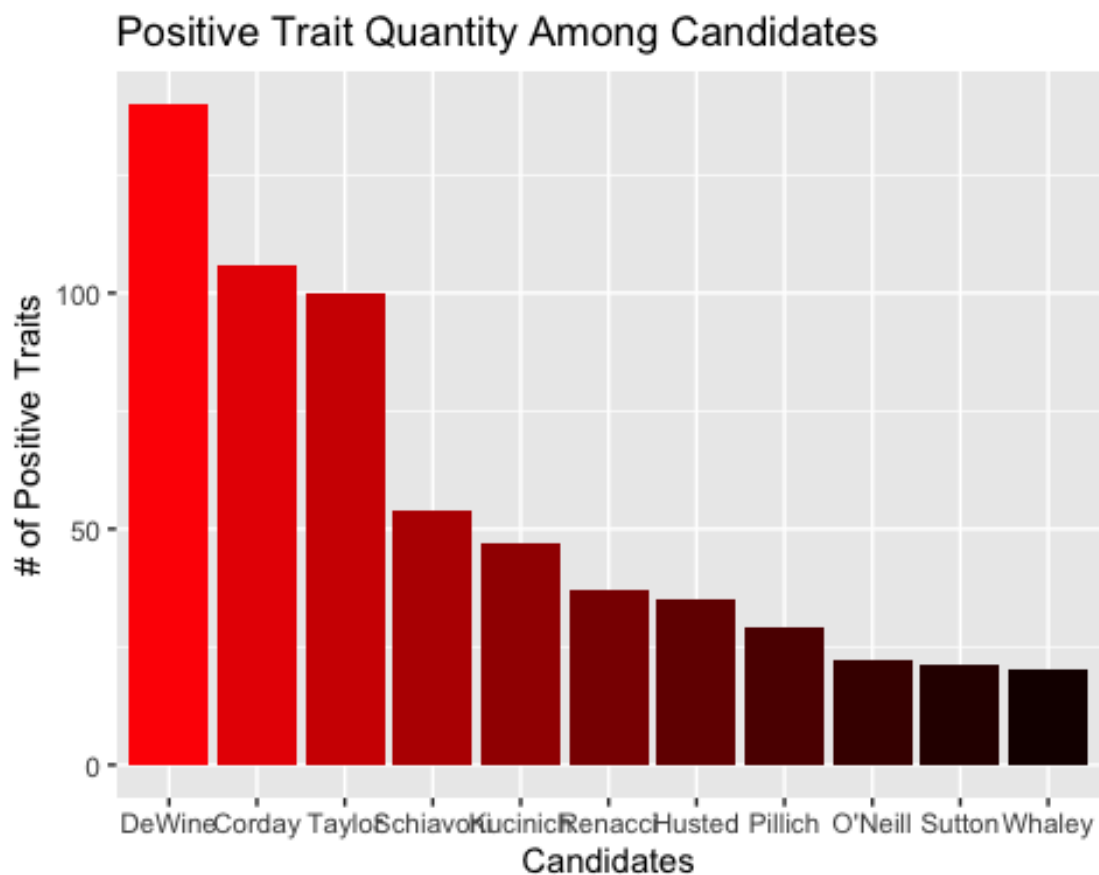
names

## [1] "Corday"      "DeWine"      "Husted"      "Kucinich"    "O'Neill"
## [6] "Pillich"     "Renacci"     "Schiavoni"   "Sutton"      "Taylor"
## [11] "Whaley"

pos <- c(106, 140, 35, 47, 22, 29, 37, 54, 21, 100, 20)
pos

## [1] 106 140 35 47 22 29 37 54 21 100 20

df3 <- data.frame(names = names, pos = pos)
ggplot(df3, aes(x=reorder(names, -pos), y= pos)) +
  geom_col(fill=rgb(11:1/11,0,0),) +labs(x = "Candidates", y = "# of Positive
Traits", title = "Positive Trait Quantity Among Candidates")
```



```
cases %>% select(MDpos,
                 MDpos1,
                 MDpos2,
                 MDpos3,
```

```

MDneg,
MDneg1,
MDneg2,
MDneg3,
RCpos,
RCpos1,
RCpos2,
RCpos3,
RCneg,
RCneg1,
RCneg2,
RCneg3,
MTpos,
MTpos1,
MTpos2,
MTpos3,
MTneg,
MTneg1,
MTneg2,
MTneg3,
) -> c2
c2$MDpos <- ifelse(c2$MDpos == "yes", 1, 0)
c2[is.na(c2)] <- 0

## Warning in `[<-.factor`(`*tmp*`, thisvar, value = 0): invalid factor
level,
## NA generated

## Warning in `[<-.factor`(`*tmp*`, thisvar, value = 0): invalid factor
level,
## NA generated

#change pos1 to 1 if it exists
c2$MDpos1 <- as.numeric(c2$MDpos1)
c2$MDpos1 <- ifelse(c2$MDpos1 == 1, 0, 1)

#change pos2 to 1 if it exists
c2$MDpos2 <- as.numeric(c2$MDpos2)
c2$MDpos2 <- ifelse(c2$MDpos2 == 1, 0, 1)

#change pos3 to 1 if it exists
c2$MDpos3 <- as.numeric(c2$MDpos3)
c2$MDpos3 <- ifelse(c2$MDpos3 == 1, 0, 1)

#Create a sum of the positive traits
c2$MDpossum <- c2$MDpos1 + c2$MDpos2 + c2$MDpos3

#Change the numbers indicating if a negative trait is mentioned at all
c2$MDneg <- ifelse(c2$MDneg == 2, 1, 0)

```

```

#change neg1 to 1 if it exists
c2$MDneg1 <- as.numeric(c2$MDneg1)
c2$MDneg1 <- ifelse(c2$MDneg1 == 1, 0, -1)

#change neg2 to 2 if it exists
c2$MDneg2 <- as.numeric(c2$MDneg2)
c2$MDneg2 <- ifelse(c2$MDneg2 == 1, 0, -1)

#change neg3 to 3 if it exists
c2$MDneg3 <- as.numeric(c2$MDneg3)
c2$MDneg3 <- ifelse(c2$MDneg3 == 1, 0, -1)

#Create a sum of all the negative traits
c2$MDnegsum <- c2$MDneg1 + c2$MDneg2 + c2$MDneg3

#Combine the positive and negative sums
c2$MDbothsum <- c2$MDpossum + c2$MDnegsum

#Cumulative sum variable
c2$MDoverallsum <- cumsum(c2$MDbothsum)

#Create a indicator variable for each article
c2$article <- c(1:510)

#Plot the article and overall sums

c2$MTpos <- ifelse(c2$MTpos == "yes", 1, 0)
c2[is.na(c2)] <- 0

## Warning in `[<-.factor`(`*tmp*`, thisvar, value = 0): invalid factor
level,
## NA generated

#change pos1 to 1 if it exists
c2$MTpos1 <- as.numeric(c2$MTpos1)
c2$MTpos1 <- ifelse(c2$MTpos1 == 1, 0, 1)

#change pos2 to 1 if it exists
c2$MTpos2 <- as.numeric(c2$MTpos2)
c2$MTpos2 <- ifelse(c2$MTpos2 == 1, 0, 1)

#change pos3 to 1 if it exists
c2$MTpos3 <- as.numeric(c2$MTpos3)
c2$MTpos3 <- ifelse(c2$MTpos3 == 1, 0, 1)

#Create a sum of the positive traits
c2$MTpossum <- c2$MTpos1 + c2$MTpos2 + c2$MTpos3

#Change the numbers indicating if a negative trait is mentioned at all

```

```

c2$MTneg <- ifelse(c2$MTneg == 2, 1, 0)

#change neg1 to 1 if it exists
c2$MTneg1 <- as.numeric(c2$MTneg1)
c2$MTneg1 <- ifelse(c2$MTneg1 == 1, 0, -1)

#change neg2 to 2 if it exists
c2$MTneg2 <- as.numeric(c2$MTneg2)
c2$MTneg2 <- ifelse(c2$MTneg2 == 1, 0, -1)

#change neg3 to 3 if it exists
c2$MTneg3 <- as.numeric(c2$MTneg3)
c2$MTneg3 <- ifelse(c2$MTneg3 == 1, 0, -1)

#Create a sum of all the negative traits
c2$MTnegsum <- c2$MTneg1 + c2$MTneg2 + c2$MTneg3

#Combine the positive and negative sums
c2$MTbothsum <- c2$MTpossum + c2$MTnegsum

#Cumulative sum variable
c2$MToverallsum <- cumsum(c2$MTbothsum)

#Create a indicator variable for each article
c2$article <- c(1:510)

#Plot the article and overall sums

c2$RCpos <- ifelse(c2$RCpos == "yes", 1, 0)
c2[is.na(c2)] <- 0

#change pos1 to 1 if it exists
c2$RCpos1 <- as.numeric(c2$RCpos1)
c2$RCpos1 <- ifelse(c2$RCpos1 == 1, 0, 1)

#change pos2 to 1 if it exists
c2$RCpos2 <- as.numeric(c2$RCpos2)
c2$RCpos2 <- ifelse(c2$RCpos2 == 1, 0, 1)

#change pos3 to 1 if it exists
c2$RCpos3 <- as.numeric(c2$RCpos3)
c2$RCpos3 <- ifelse(c2$RCpos3 == 1, 0, 1)

#Create a sum of the positive traits
c2$RCpossum <- c2$RCpos1 + c2$RCpos2 + c2$RCpos3

#Change the numbers indicating if a negative trait is mentioned at all
c2$RCneg <- ifelse(c2$RCneg == 2, 1, 0)

```

```

#change neg1 to 1 if it exists
c2$RCneg1 <- as.numeric(c2$RCneg1)
c2$RCneg1 <- ifelse(c2$RCneg1 == 1, 0, -1)

#change neg2 to 2 if it exists
c2$RCneg2 <- as.numeric(c2$RCneg2)
c2$RCneg2 <- ifelse(c2$RCneg2 == 1, 0, -1)

#change neg3 to 3 if it exists
c2$RCneg3 <- as.numeric(c2$RCneg3)
c2$RCneg3 <- ifelse(c2$RCneg3 == 1, 0, -1)

#Create a sum of all the negative traits
c2$RCnegsum <- c2$RCneg1 + c2$RCneg2 + c2$RCneg3

#Combine the positive and negative sums
c2$RCbothsum <- c2$RCpossum + c2$RCnegsum

#Cumulative sum variable
c2$RCoverallsum <- cumsum(c2$RCbothsum)

#Create a indicator variable for each article
c2$article <- c(1:510)

#Plot the article and overall sums
ggplot(c2) + geom_line(aes(article, MDoverallsum), col = "red") +
geom_line(aes(article, RCoverallsum), color = 'blue') +
geom_line(aes(article, MToverallsum), col = "green")+ labs(x = "Article", y =
"Quality Points", title = "Candidate Overall Quality Points") + theme(
  legend.position = c(0.9, 0.9),
  legend.justification = c("left", "top"))

```

Candidate Overall Quality Points

