

Data 607 Homework 1

Jean Jimenez

2023-09-01

Overview

The article I chose was "Trump's Endorsees Have Started Losing More. But Don't Read Into That For 2024."

This article is an analysis of the 2022 election data and gives us a glimpse into how effective Trump's endorsements are. While Trump boasts that his endorsements are a guaranteed win, the election candidate tells a different story. Trump's endorsement has an overall success rate of 95%. However, there are reasons other than Trump endorsement that had influenced this number. 32% of Trump's endorsements went to candidates that had no opposition when running. Also, 44% of all of Trump's endorsements were incumbents. These two factors boost Trump's endorsement success rate and allow him to misinform his followers. In reality, his endorsement's success rate is a lower 82% overall. The author also analyzes the statistics in different states and saw differences in success rates based on many different factors. I will work with the data provided to modify it and see if I can come up with any conclusions for NY Republican Candidates that ran in 2022.

Cleaning the Data

First I imported the data set from the linked GITHUB.

```
#importing the data set

library(readr)

rep_candidates=read.csv(url("https://raw.githubusercontent.com/fivethirtyeight/data/master/primary-proj
head(rep_candidates)
```

##	Candidate	Gender	Race.1	Race.2	Race.3	Incumbent
## 1	Aditya "A.D." Atholi	Male	Asian	(Indian)		No
## 2	Joe McDaniel	Male	White			No
## 3	Nathaniel Moran	Male	White			No
## 4	John Porro	Male	White			No
## 5	Dan Crenshaw	Male	White			Yes
## 6	Jameson Ellis	Male	White			No
##	Incumbent.Challenger	State	Primary.Date	Office	District	Primary.Votes
## 1	No	Texas	3/1/22	Representative	1	6,186
## 2	No	Texas	3/1/22	Representative	1	19,708
## 3	No	Texas	3/1/22	Representative	1	51,312
## 4	No	Texas	3/1/22	Representative	1	4,238

## 5	No Texas	3/1/22	Representative	2	45,863
## 6	Yes Texas	3/1/22	Representative	2	10,195
##	Primary.. Primary.Outcome	Runoff.Votes	Runoff..	Runoff.Outcome	
## 1	8%	Lost	N/A	N/A	N/A
## 2	24%	Lost	N/A	N/A	N/A
## 3	63%	Won	N/A	N/A	N/A
## 4	5%	Lost	N/A	N/A	N/A
## 5	74%	Won	N/A	N/A	N/A
## 6	17%	Lost	N/A	N/A	N/A
##	X2020.Election.Stance	Trump	Trump.Date	Club.for.Growth	Party.Committee
## 1	No comment	N/A	N/A	N/A	N/A
## 2	Raised questions	N/A	N/A	N/A	N/A
## 3	Raised questions	N/A	N/A	N/A	N/A
## 4	No comment	N/A	N/A	N/A	N/A
## 5	Accepted with reservations	N/A	N/A	N/A	N/A
## 6	Fully denied	N/A	N/A	N/A	N/A
##	Renew.America	E.PAC	VIEW.PAC	Maggie.s.List	Winning.for.Women
## 1	N/A	N/A	N/A	N/A	N/A
## 2	N/A	N/A	N/A	N/A	N/A
## 3	N/A	N/A	N/A	N/A	N/A
## 4	N/A	N/A	N/A	N/A	N/A
## 5	N/A	N/A	N/A	N/A	N/A
## 6	N/A	N/A	N/A	N/A	N/A

```
dim(rep_candidates)
```

```
## [1] 1599 27
```

```
#rep_candidates is a 1599 x 27 dataframe
```

rep_candidates is a data frame that is 1599 x 27.

I am only interested in looking at NY Republican candidates.

I will narrow it down using dplyr, and filter out only rows with New York under State

```
#Getting a subset of the data that we are interested in (NY)  
#using dplyr because its easier to filter rows with it in my opinion
```

```
library(dplyr)
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
## filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
## intersect, setdiff, setequal, union
```

```
ny_rep_candidates = rep_candidates %>% filter(State == 'New York')
dim(ny_rep_candidates)
```

```
## [1] 43 27
```

Now, our new data frame ny_rep_candidates has 43 rows and 27 columns.

There are 43 Republican Candidates that ran in NY in 2020.

I'm not interested in all of the columns so now I'm going to remove some of them.

```
#There are 43 Republican Candidates running in new york
```

```
#Getting rid of columns that we dont need
```

```
ny_rep_candidates$Race.2=NULL
ny_rep_candidates$Race.3=NULL
ny_rep_candidates$Primary.Date=NULL
ny_rep_candidates$Club.for.Growth=NULL
ny_rep_candidates$Party.Committee=NULL
ny_rep_candidates$Renew.America=NULL
ny_rep_candidates$VIEW.PAC=NULL
ny_rep_candidates$Maggie.s.List=NULL
ny_rep_candidates$Winning.for.Women=NULL
ny_rep_candidates$E.PAC=NULL
ny_rep_candidates$Candidate=NULL
ny_rep_candidates$Trump.Date=NULL
ny_rep_candidates
```

##	Gender	Race.1	Incumbent	Incumbent.Challenger	State
## 1	Male	Black	No	No	New York
## 2	Male	White	No	No	New York
## 3	Male	White	No	No	New York
## 4	Male	White	No	No	New York
## 5	Male	White	No	No	New York
## 6	Male	White	No	No	New York
## 7	Male	White	No	No	New York
## 8	Female	White	No	No	New York
## 9	Male	White	Yes	No	New York
## 10	Male	White	No	Yes	New York
## 11	Male	White	No	Yes	New York
## 12	Male	White	No	No	New York
## 13	Male	White	No	No	New York
## 14	Male	White	No	No	New York
## 15	Male	White	No	No	New York
## 16	Male	Latino	No	No	New York
## 17	Male	White	No	No	New York
## 18	Female	Middle Eastern (Lebanese)	No	No	New York
## 19	Female	White	Yes	No	New York
## 20	Male	White	No	Yes	New York
## 21	Male	White	No	No	New York
## 22	Female	White	No	No	New York
## 23	Male	Latino (Cuban)	No	No	New York

## 24	Male		White	No		No New York
## 25	Female		White	No		No New York
## 26	Male		White	No		No New York
## 27	Male		White	No		No New York
## 28	Male		White	No		No New York
## 29	Female		Unknown	No		No New York
## 30	Male		White	No		No New York
## 31	Male		White	No		No New York
## 32	Male		White	No		No New York
## 33	Female		White	No		No New York
## 34	Female		White	Yes		No New York
## 35	Male		White	No		No New York
## 36	Male		White	No		No New York
## 37	Male		White	No		No New York
## 38	Male		White	No		No New York
## 39	Female		White	Yes		No New York
## 40	Male		White	No		Yes New York
## 41	Male		White	No		Yes New York
## 42	Male		Black	No		No New York
## 43	Male		White	No		No New York
##	Office	District	Primary.Votes	Primary..	Primary.Outcome	Runoff.Votes
## 1	Senator	N/A	1	100%	Won	N/A
## 2	Governor	N/A	103,267	23%	Lost	N/A
## 3	Governor	N/A	66,736	15%	Lost	N/A
## 4	Governor	N/A	84,464	19%	Lost	N/A
## 5	Governor	N/A	196,874	43%	Won	N/A
## 6	Representative	1	6,391	25%	Lost	N/A
## 7	Representative	1	12,015	47%	Won	N/A
## 8	Representative	1	7,015	28%	Lost	N/A
## 9	Representative	2	9,902	53%	Won	N/A
## 10	Representative	2	7,250	38%	Lost	N/A
## 11	Representative	2	1,622	9%	Lost	N/A
## 12	Representative	3	1	100%	Won	N/A
## 13	Representative	4	1	100%	Won	N/A
## 14	Representative	5	1	100%	Won	N/A
## 15	Representative	6	1	100%	Won	N/A
## 16	Representative	7	1	100%	Won	N/A
## 17	Representative	8	1	100%	Won	N/A
## 18	Representative	10	1	100%	Won	N/A
## 19	Representative	11	12,431	78%	Won	N/A
## 20	Representative	11	3,407	21%	Lost	N/A
## 21	Representative	12	1	100%	Won	N/A
## 22	Representative	14	1,608	67%	Won	N/A
## 23	Representative	14	761	32%	Lost	N/A
## 24	Representative	15	1	100%	Won	N/A
## 25	Representative	16	1	100%	Won	N/A
## 26	Representative	17	1,392	8%	Lost	N/A
## 27	Representative	17	12,317	75%	Won	N/A
## 28	Representative	17	188	1%	Lost	N/A
## 29	Representative	17	491	3%	Lost	N/A
## 30	Representative	17	1,958	12%	Lost	N/A
## 31	Representative	18	1	100%	Won	N/A
## 32	Representative	19	1	100%	Won	N/A
## 33	Representative	20	1	100%	Won	N/A

## 34 Representative	21	1	100%	Won	N/A
## 35 Representative	22	14,351	57%	Won	N/A
## 36 Representative	22	10,501	42%	Lost	N/A
## 37 Representative	23	22,603	47%	Lost	N/A
## 38 Representative	23	24,450	51%	Won	N/A
## 39 Representative	24	17,630	54%	Won	N/A
## 40 Representative	24	13,150	40%	Lost	N/A
## 41 Representative	24	1,967	6%	Lost	N/A
## 42 Representative	25	1	100%	Won	N/A
## 43 Representative	26	1	100%	Won	N/A
##	Runoff..	Runoff.Outcome	X2020.Election.Stance	Trump	
## 1	N/A	N/A	Fully accepted	N/A	
## 2	N/A	N/A	Fully denied	N/A	
## 3	N/A	N/A	Fully accepted	N/A	
## 4	N/A	N/A	No comment	N/A	
## 5	N/A	N/A	Accepted with reservations	N/A	
## 6	N/A	N/A	Fully accepted	N/A	
## 7	N/A	N/A	Raised questions	N/A	
## 8	N/A	N/A	Raised questions	N/A	
## 9	N/A	N/A	Fully accepted	N/A	
## 10	N/A	N/A	Fully denied	N/A	
## 11	N/A	N/A	Fully denied	N/A	
## 12	N/A	N/A	Fully denied	N/A	
## 13	N/A	N/A	Fully accepted	N/A	
## 14	N/A	N/A	Accepted with reservations	N/A	
## 15	N/A	N/A	Raised questions	N/A	
## 16	N/A	N/A	Fully accepted	N/A	
## 17	N/A	N/A	No comment	N/A	
## 18	N/A	N/A	Raised questions	N/A	
## 19	N/A	N/A	Fully denied	Yes	
## 20	N/A	N/A	Raised questions	No	
## 21	N/A	N/A	Accepted with reservations	N/A	
## 22	N/A	N/A	Fully denied	N/A	
## 23	N/A	N/A	No comment	N/A	
## 24	N/A	N/A	No comment	N/A	
## 25	N/A	N/A	Avoided answering	N/A	
## 26	N/A	N/A	No comment	N/A	
## 27	N/A	N/A	Fully accepted	N/A	
## 28	N/A	N/A	No comment	N/A	
## 29	N/A	N/A	No comment	N/A	
## 30	N/A	N/A	No comment	N/A	
## 31	N/A	N/A	Avoided answering	N/A	
## 32	N/A	N/A	No comment	N/A	
## 33	N/A	N/A	Fully denied	N/A	
## 34	N/A	N/A	Fully denied	Yes	
## 35	N/A	N/A	Accepted with reservations	N/A	
## 36	N/A	N/A	No comment	N/A	
## 37	N/A	N/A	No comment	N/A	
## 38	N/A	N/A	No comment	N/A	
## 39	N/A	N/A	Raised questions	Yes	
## 40	N/A	N/A	Fully denied	No	
## 41	N/A	N/A	No comment	No	
## 42	N/A	N/A	No comment	N/A	
## 43	N/A	N/A	No comment	N/A	

```
dim(ny_rep_candidates)
```

```
## [1] 43 15
```

```
#Now we have 43 rows and 17 columns
```

Now we have 17 columns.

I will rename all the columns because they are currently too long.

```
#Now lets rename all of the column into more useful names
```

```
head(ny_rep_candidates)
```

```
##   Gender Race.1 Incumbent Incumbent.Challenger      State      Office District
## 1   Male  Black        No                No New York      Senator      N/A
## 2   Male  White        No                No New York      Governor     N/A
## 3   Male  White        No                No New York      Governor     N/A
## 4   Male  White        No                No New York      Governor     N/A
## 5   Male  White        No                No New York      Governor     N/A
## 6   Male  White        No                No New York Representative 1
##   Primary.Votes Primary.. Primary.Outcome Runoff.Votes Runoff.. Runoff.Outcome
## 1              1      100%           Won           N/A      N/A           N/A
## 2          103,267      23%          Lost           N/A      N/A           N/A
## 3           66,736      15%          Lost           N/A      N/A           N/A
## 4           84,464      19%          Lost           N/A      N/A           N/A
## 5          196,874      43%           Won           N/A      N/A           N/A
## 6           6,391      25%          Lost           N/A      N/A           N/A
##   X2020.Election.Stance Trump
## 1              Fully accepted  N/A
## 2              Fully denied  N/A
## 3              Fully accepted  N/A
## 4              No comment    N/A
## 5 Accepted with reservations  N/A
## 6              Fully accepted  N/A
```

```
colnames(ny_rep_candidates) <- c("gender", "race", "inc", "inc chal", "st", "pos", "dist", "n_prim", "rat_prim",
                                "n_run", "rat_run", "out_run", "insurrectionist", "trump_endorse")
```

```
head(ny_rep_candidates)
```

```
##   gender race inc inc chal      st      pos dist  n_prim rat_prim
## 1   Male Black No      No New York      Senator  N/A      1    100%
## 2   Male White No      No New York      Governor  N/A 103,267    23%
## 3   Male White No      No New York      Governor  N/A  66,736    15%
## 4   Male White No      No New York      Governor  N/A  84,464    19%
## 5   Male White No      No New York      Governor  N/A 196,874    43%
## 6   Male White No      No New York Representative 1   6,391    25%
##   out_prim n_run rat_run out_run      insurrectionist trump_endorse
## 1      Won  N/A    N/A    N/A      Fully accepted      N/A
## 2     Lost  N/A    N/A    N/A      Fully denied      N/A
## 3     Lost  N/A    N/A    N/A      Fully accepted      N/A
## 4     Lost  N/A    N/A    N/A      No comment      N/A
## 5      Won  N/A    N/A    N/A Accepted with reservations  N/A
## 6     Lost  N/A    N/A    N/A      Fully accepted      N/A
```

Now I will make the insurrectionist column either True or False. If the candidate denied the results of the 2020 Presidential election, it will be set to True. To do this, I looked for all the unique possibilities in this column and replaced it using an if statement.

```
#Now will replace the values in the col insurrectionist
#Column now is going to be boolean
```

```
unique(ny_rep_candidates$insurrectionist)
```

```
## [1] "Fully accepted"      "Fully denied"
## [3] "No comment"         "Accepted with reservations"
## [5] "Raised questions"    "Avoided answering"
```

```
if (any(ny_rep_candidates$insurrectionist == "Fully accepted")) {
  ny_rep_candidates$insurrectionist=FALSE
} else {
  ny_rep_candidates$insurrectionist = TRUE
}
```

```
ny_rep_candidates
```

##	gender	race	inc	inc	chal	st	pos	dist
## 1	Male	Black	No		No	New York	Senator	N/A
## 2	Male	White	No		No	New York	Governor	N/A
## 3	Male	White	No		No	New York	Governor	N/A
## 4	Male	White	No		No	New York	Governor	N/A
## 5	Male	White	No		No	New York	Governor	N/A
## 6	Male	White	No		No	New York	Representative	1
## 7	Male	White	No		No	New York	Representative	1
## 8	Female	White	No		No	New York	Representative	1
## 9	Male	White	Yes		No	New York	Representative	2
## 10	Male	White	No		Yes	New York	Representative	2
## 11	Male	White	No		Yes	New York	Representative	2
## 12	Male	White	No		No	New York	Representative	3
## 13	Male	White	No		No	New York	Representative	4
## 14	Male	White	No		No	New York	Representative	5
## 15	Male	White	No		No	New York	Representative	6
## 16	Male	Latino	No		No	New York	Representative	7
## 17	Male	White	No		No	New York	Representative	8
## 18	Female	Middle Eastern (Lebanese)	No		No	New York	Representative	10
## 19	Female	White	Yes		No	New York	Representative	11
## 20	Male	White	No		Yes	New York	Representative	11
## 21	Male	White	No		No	New York	Representative	12
## 22	Female	White	No		No	New York	Representative	14
## 23	Male	Latino (Cuban)	No		No	New York	Representative	14
## 24	Male	White	No		No	New York	Representative	15
## 25	Female	White	No		No	New York	Representative	16
## 26	Male	White	No		No	New York	Representative	17
## 27	Male	White	No		No	New York	Representative	17
## 28	Male	White	No		No	New York	Representative	17
## 29	Female	Unknown	No		No	New York	Representative	17
## 30	Male	White	No		No	New York	Representative	17
## 31	Male	White	No		No	New York	Representative	18

## 32	Male	White	No	No New York Representative	19
## 33	Female	White	No	No New York Representative	20
## 34	Female	White	Yes	No New York Representative	21
## 35	Male	White	No	No New York Representative	22
## 36	Male	White	No	No New York Representative	22
## 37	Male	White	No	No New York Representative	23
## 38	Male	White	No	No New York Representative	23
## 39	Female	White	Yes	No New York Representative	24
## 40	Male	White	No	Yes New York Representative	24
## 41	Male	White	No	Yes New York Representative	24
## 42	Male	Black	No	No New York Representative	25
## 43	Male	White	No	No New York Representative	26

##	n_prim	rat_prim	out_prim	n_run	rat_run	out_run	insurrectionist
## 1	1	100%	Won	N/A	N/A	N/A	FALSE
## 2	103,267	23%	Lost	N/A	N/A	N/A	FALSE
## 3	66,736	15%	Lost	N/A	N/A	N/A	FALSE
## 4	84,464	19%	Lost	N/A	N/A	N/A	FALSE
## 5	196,874	43%	Won	N/A	N/A	N/A	FALSE
## 6	6,391	25%	Lost	N/A	N/A	N/A	FALSE
## 7	12,015	47%	Won	N/A	N/A	N/A	FALSE
## 8	7,015	28%	Lost	N/A	N/A	N/A	FALSE
## 9	9,902	53%	Won	N/A	N/A	N/A	FALSE
## 10	7,250	38%	Lost	N/A	N/A	N/A	FALSE
## 11	1,622	9%	Lost	N/A	N/A	N/A	FALSE
## 12	1	100%	Won	N/A	N/A	N/A	FALSE
## 13	1	100%	Won	N/A	N/A	N/A	FALSE
## 14	1	100%	Won	N/A	N/A	N/A	FALSE
## 15	1	100%	Won	N/A	N/A	N/A	FALSE
## 16	1	100%	Won	N/A	N/A	N/A	FALSE
## 17	1	100%	Won	N/A	N/A	N/A	FALSE
## 18	1	100%	Won	N/A	N/A	N/A	FALSE
## 19	12,431	78%	Won	N/A	N/A	N/A	FALSE
## 20	3,407	21%	Lost	N/A	N/A	N/A	FALSE
## 21	1	100%	Won	N/A	N/A	N/A	FALSE
## 22	1,608	67%	Won	N/A	N/A	N/A	FALSE
## 23	761	32%	Lost	N/A	N/A	N/A	FALSE
## 24	1	100%	Won	N/A	N/A	N/A	FALSE
## 25	1	100%	Won	N/A	N/A	N/A	FALSE
## 26	1,392	8%	Lost	N/A	N/A	N/A	FALSE
## 27	12,317	75%	Won	N/A	N/A	N/A	FALSE
## 28	188	1%	Lost	N/A	N/A	N/A	FALSE
## 29	491	3%	Lost	N/A	N/A	N/A	FALSE
## 30	1,958	12%	Lost	N/A	N/A	N/A	FALSE
## 31	1	100%	Won	N/A	N/A	N/A	FALSE
## 32	1	100%	Won	N/A	N/A	N/A	FALSE
## 33	1	100%	Won	N/A	N/A	N/A	FALSE
## 34	1	100%	Won	N/A	N/A	N/A	FALSE
## 35	14,351	57%	Won	N/A	N/A	N/A	FALSE
## 36	10,501	42%	Lost	N/A	N/A	N/A	FALSE
## 37	22,603	47%	Lost	N/A	N/A	N/A	FALSE
## 38	24,450	51%	Won	N/A	N/A	N/A	FALSE
## 39	17,630	54%	Won	N/A	N/A	N/A	FALSE
## 40	13,150	40%	Lost	N/A	N/A	N/A	FALSE
## 41	1,967	6%	Lost	N/A	N/A	N/A	FALSE

##	42	1	100%	Won	N/A	N/A	N/A	FALSE
##	43	1	100%	Won	N/A	N/A	N/A	FALSE
##	trump_endorse							
##	1		N/A					
##	2		N/A					
##	3		N/A					
##	4		N/A					
##	5		N/A					
##	6		N/A					
##	7		N/A					
##	8		N/A					
##	9		N/A					
##	10		N/A					
##	11		N/A					
##	12		N/A					
##	13		N/A					
##	14		N/A					
##	15		N/A					
##	16		N/A					
##	17		N/A					
##	18		N/A					
##	19		Yes					
##	20		No					
##	21		N/A					
##	22		N/A					
##	23		N/A					
##	24		N/A					
##	25		N/A					
##	26		N/A					
##	27		N/A					
##	28		N/A					
##	29		N/A					
##	30		N/A					
##	31		N/A					
##	32		N/A					
##	33		N/A					
##	34		Yes					
##	35		N/A					
##	36		N/A					
##	37		N/A					
##	38		N/A					
##	39		Yes					
##	40		No					
##	41		No					
##	42		N/A					
##	43		N/A					

Now, I will filter out so that only candidates that won are in our data frame.

I will also remove some other columns that we don't need anymore.

```
#We only care about the winners. Filtering only winners.
```

```
ny_rep_candidates = ny_rep_candidates %>% filter(out_prim == 'Won')
```

#now some columns are useless so removing

```
ny_rep_candidates$st=NULL
ny_rep_candidates$out_prim=NULL
ny_rep_candidates$n_run=NULL
ny_rep_candidates$rat_run=NULL
ny_rep_candidates$out_run=NULL
ny_rep_candidates
```

##	gender	race	inc	inc	chal	pos	dist	n_prim
## 1	Male	Black	No		No	Senator	N/A	1
## 2	Male	White	No		No	Governor	N/A	196,874
## 3	Male	White	No		No Representative		1	12,015
## 4	Male	White	Yes		No Representative		2	9,902
## 5	Male	White	No		No Representative		3	1
## 6	Male	White	No		No Representative		4	1
## 7	Male	White	No		No Representative		5	1
## 8	Male	White	No		No Representative		6	1
## 9	Male	Latino	No		No Representative		7	1
## 10	Male	White	No		No Representative		8	1
## 11	Female Middle Eastern (Lebanese)		No		No Representative		10	1
## 12	Female	White	Yes		No Representative		11	12,431
## 13	Male	White	No		No Representative		12	1
## 14	Female	White	No		No Representative		14	1,608
## 15	Male	White	No		No Representative		15	1
## 16	Female	White	No		No Representative		16	1
## 17	Male	White	No		No Representative		17	12,317
## 18	Male	White	No		No Representative		18	1
## 19	Male	White	No		No Representative		19	1
## 20	Female	White	No		No Representative		20	1
## 21	Female	White	Yes		No Representative		21	1
## 22	Male	White	No		No Representative		22	14,351
## 23	Male	White	No		No Representative		23	24,450
## 24	Female	White	Yes		No Representative		24	17,630
## 25	Male	Black	No		No Representative		25	1
## 26	Male	White	No		No Representative		26	1
##	rat_prim	insurrectionist	trump_endorse					
## 1	100%	FALSE	N/A					
## 2	43%	FALSE	N/A					
## 3	47%	FALSE	N/A					
## 4	53%	FALSE	N/A					
## 5	100%	FALSE	N/A					
## 6	100%	FALSE	N/A					
## 7	100%	FALSE	N/A					
## 8	100%	FALSE	N/A					
## 9	100%	FALSE	N/A					
## 10	100%	FALSE	N/A					
## 11	100%	FALSE	N/A					
## 12	78%	FALSE	Yes					
## 13	100%	FALSE	N/A					
## 14	67%	FALSE	N/A					
## 15	100%	FALSE	N/A					
## 16	100%	FALSE	N/A					
## 17	75%	FALSE	N/A					

## 18	100%	FALSE	N/A
## 19	100%	FALSE	N/A
## 20	100%	FALSE	N/A
## 21	100%	FALSE	Yes
## 22	57%	FALSE	N/A
## 23	51%	FALSE	N/A
## 24	54%	FALSE	Yes
## 25	100%	FALSE	N/A
## 26	100%	FALSE	N/A

Replacing values under a column. True if Trump endorsed the candidate and false if he didn't.

```
#Now will replace the values in the trump endorse
#To True or false

if (any(ny_rep_candidates$trump_endorse == "Yes")) {
  ny_rep_candidates$trump_endorse=TRUE
} else {
  ny_rep_candidates$trump_endorse = FALSE
}
```

Data Exploration and Visualization

Now I will play around with the data a little bit.

Pie chart of the racial distribution of republican candidates in NY that won:

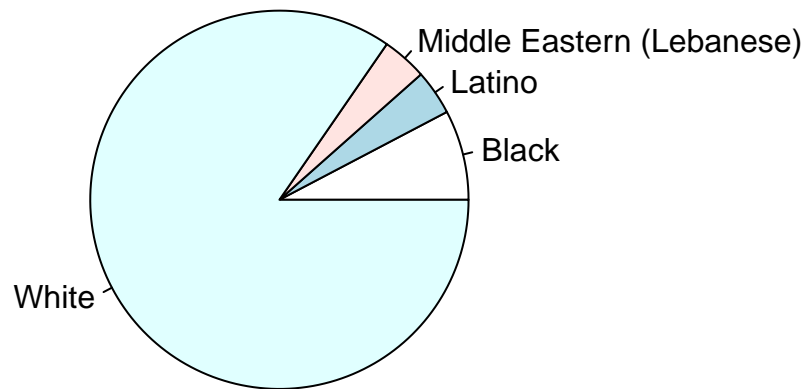
```
#now lets analyze and visualize
ny_rep_candidates
```

##	gender	race	inc	inc	chal	pos	dist	n_prim
## 1	Male	Black	No		No	Senator	N/A	1
## 2	Male	White	No		No	Governor	N/A	196,874
## 3	Male	White	No		No	Representative	1	12,015
## 4	Male	White	Yes		No	Representative	2	9,902
## 5	Male	White	No		No	Representative	3	1
## 6	Male	White	No		No	Representative	4	1
## 7	Male	White	No		No	Representative	5	1
## 8	Male	White	No		No	Representative	6	1
## 9	Male	Latino	No		No	Representative	7	1
## 10	Male	White	No		No	Representative	8	1
## 11	Female	Middle Eastern (Lebanese)	No		No	Representative	10	1
## 12	Female	White	Yes		No	Representative	11	12,431
## 13	Male	White	No		No	Representative	12	1
## 14	Female	White	No		No	Representative	14	1,608
## 15	Male	White	No		No	Representative	15	1
## 16	Female	White	No		No	Representative	16	1
## 17	Male	White	No		No	Representative	17	12,317
## 18	Male	White	No		No	Representative	18	1
## 19	Male	White	No		No	Representative	19	1
## 20	Female	White	No		No	Representative	20	1
## 21	Female	White	Yes		No	Representative	21	1
## 22	Male	White	No		No	Representative	22	14,351

## 23	Male	White	No	No Representative	23	24,450
## 24	Female	White	Yes	No Representative	24	17,630
## 25	Male	Black	No	No Representative	25	1
## 26	Male	White	No	No Representative	26	1
##	rat_prim	insurrectionist	trump_endorse			
## 1	100%	FALSE	TRUE			
## 2	43%	FALSE	TRUE			
## 3	47%	FALSE	TRUE			
## 4	53%	FALSE	TRUE			
## 5	100%	FALSE	TRUE			
## 6	100%	FALSE	TRUE			
## 7	100%	FALSE	TRUE			
## 8	100%	FALSE	TRUE			
## 9	100%	FALSE	TRUE			
## 10	100%	FALSE	TRUE			
## 11	100%	FALSE	TRUE			
## 12	78%	FALSE	TRUE			
## 13	100%	FALSE	TRUE			
## 14	67%	FALSE	TRUE			
## 15	100%	FALSE	TRUE			
## 16	100%	FALSE	TRUE			
## 17	75%	FALSE	TRUE			
## 18	100%	FALSE	TRUE			
## 19	100%	FALSE	TRUE			
## 20	100%	FALSE	TRUE			
## 21	100%	FALSE	TRUE			
## 22	57%	FALSE	TRUE			
## 23	51%	FALSE	TRUE			
## 24	54%	FALSE	TRUE			
## 25	100%	FALSE	TRUE			
## 26	100%	FALSE	TRUE			

#race distribution of NY Republican candidates that won in 2020

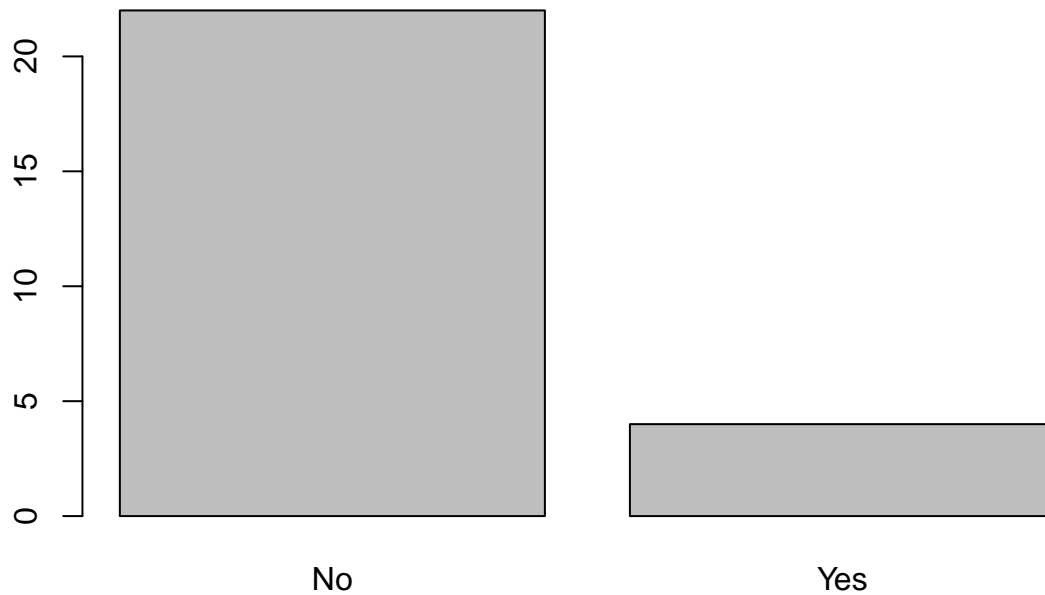
```
pie(table(ny_rep_candidates$race))
```



Barplot showing candidate incumbent count

```
#Was Candidate incumbent?
```

```
barplot(table(ny_rep_candidates$inc))
```



I want to measure the mean ratio of votes. For that, I have to first remove the % symbol from the value because it is a character and not a numeric.

```
without_percent=c()

for(i in ny_rep_candidates$rat_prim){
  i=substr(i, start=1, stop=nchar(i)-1)
  without_percent=append(without_percent,i)
}
without_percent
```

```
## [1] "100" "43"  "47"  "53"  "100" "100" "100" "100" "100" "100" "100" "78"
## [13] "100" "67"  "100" "100" "75"  "100" "100" "100" "100" "57"  "51"  "54"
## [25] "100" "100"
```

Now that the % sign is removed, we can convert it back to numeric and make it into an actual ratio.

```
ny_rep_candidates$rat_prim=as.numeric(without_percent)

#Now lets make this an actual ratio

ny_rep_candidates$rat_prim=(ny_rep_candidates$rat_prim)/100
ny_rep_candidates
```

```
##      gender      race inc inc chal      pos dist  n_prim
```

## 1	Male	Black	No	No	Senator	N/A	1
## 2	Male	White	No	No	Governor	N/A	196,874
## 3	Male	White	No	No	Representative	1	12,015
## 4	Male	White	Yes	No	Representative	2	9,902
## 5	Male	White	No	No	Representative	3	1
## 6	Male	White	No	No	Representative	4	1
## 7	Male	White	No	No	Representative	5	1
## 8	Male	White	No	No	Representative	6	1
## 9	Male	Latino	No	No	Representative	7	1
## 10	Male	White	No	No	Representative	8	1
## 11	Female	Middle Eastern (Lebanese)	No	No	Representative	10	1
## 12	Female	White	Yes	No	Representative	11	12,431
## 13	Male	White	No	No	Representative	12	1
## 14	Female	White	No	No	Representative	14	1,608
## 15	Male	White	No	No	Representative	15	1
## 16	Female	White	No	No	Representative	16	1
## 17	Male	White	No	No	Representative	17	12,317
## 18	Male	White	No	No	Representative	18	1
## 19	Male	White	No	No	Representative	19	1
## 20	Female	White	No	No	Representative	20	1
## 21	Female	White	Yes	No	Representative	21	1
## 22	Male	White	No	No	Representative	22	14,351
## 23	Male	White	No	No	Representative	23	24,450
## 24	Female	White	Yes	No	Representative	24	17,630
## 25	Male	Black	No	No	Representative	25	1
## 26	Male	White	No	No	Representative	26	1
##	rat_prim	insurrectionist	trump_endorse				
## 1	1.00	FALSE	TRUE				
## 2	0.43	FALSE	TRUE				
## 3	0.47	FALSE	TRUE				
## 4	0.53	FALSE	TRUE				
## 5	1.00	FALSE	TRUE				
## 6	1.00	FALSE	TRUE				
## 7	1.00	FALSE	TRUE				
## 8	1.00	FALSE	TRUE				
## 9	1.00	FALSE	TRUE				
## 10	1.00	FALSE	TRUE				
## 11	1.00	FALSE	TRUE				
## 12	0.78	FALSE	TRUE				
## 13	1.00	FALSE	TRUE				
## 14	0.67	FALSE	TRUE				
## 15	1.00	FALSE	TRUE				
## 16	1.00	FALSE	TRUE				
## 17	0.75	FALSE	TRUE				
## 18	1.00	FALSE	TRUE				
## 19	1.00	FALSE	TRUE				
## 20	1.00	FALSE	TRUE				
## 21	1.00	FALSE	TRUE				
## 22	0.57	FALSE	TRUE				
## 23	0.51	FALSE	TRUE				
## 24	0.54	FALSE	TRUE				
## 25	1.00	FALSE	TRUE				
## 26	1.00	FALSE	TRUE				

```
#finally, lets measure the mean ratio of votes one during the primary  
mean(ny_rep_candidates$rat_prim)
```

```
## [1] 0.8557692
```

NY Republican candidates won 85% of the primary vote in 2022.

Findings and Recommendations

NY Republican candidates won 85% of the primary vote in 2022.

This is a really robust data set. The 2022 election cycle was an interesting one. Right wing media and polls boasted about a red wave that was supposed to take over the house and senate that season. In reality, the results were complete opposite of what was expected.

There was an unexpected red wave in New York State, where a lot of republicans won seats in districts that weren't expected to be republican. This is likely an effect of the change in electoral map from the previous season to the current. The current electoral map is gerrymandered to favor republicans. It would be interesting in the future to compare these results with the results of the previous election.

It would also be interesting to compare NY results vs results from other states to see if there were any differences.