Basic Loading and Transformations

David Moste

2/1/2020

Introduction

Getting the Data

I started by loading the .csv files into two data frames.

```
cable_data <- data.frame(read.csv("media-mentions-2020/cable_weekly.csv"))
online_data <- data.frame(read.csv("media-mentions-2020/online_weekly.csv"))
```

The next step was to combine the relevant columns into a new data frame titled all_data and provide relevant column names.

```
all_data <- data.frame(cable_data$date, cable_data$name,
                       cable_data$pct_of_all_candidate_clips,
                       online_data$pct_of_all_candidate_stories)
names(all_data) <- c("Date","Name","Clips","Stories")
```

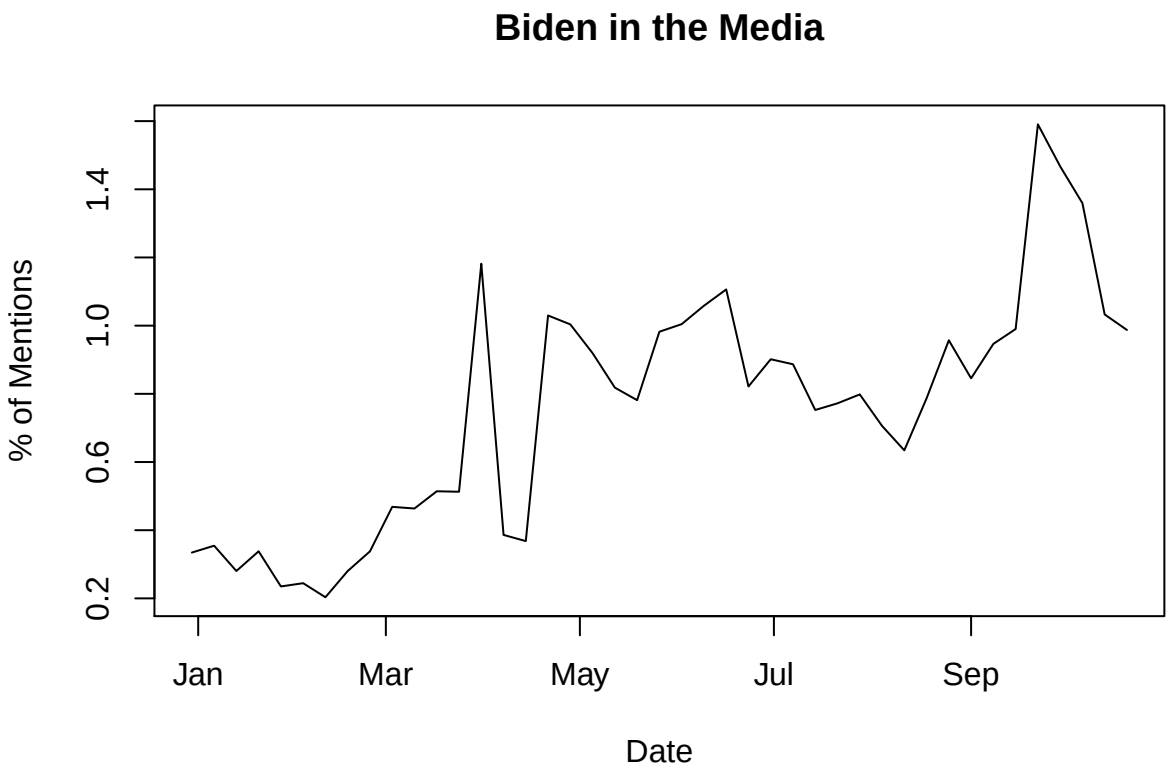
Finally, I subsetted the data to only include mentions of Joe Biden. At the same time, I converted the Date column into Date objects.

```
all_biden <- subset(all_data, Name == "Joe Biden")
all_biden$Date <- as.Date(all_biden$Date)
```

Looking at the Data

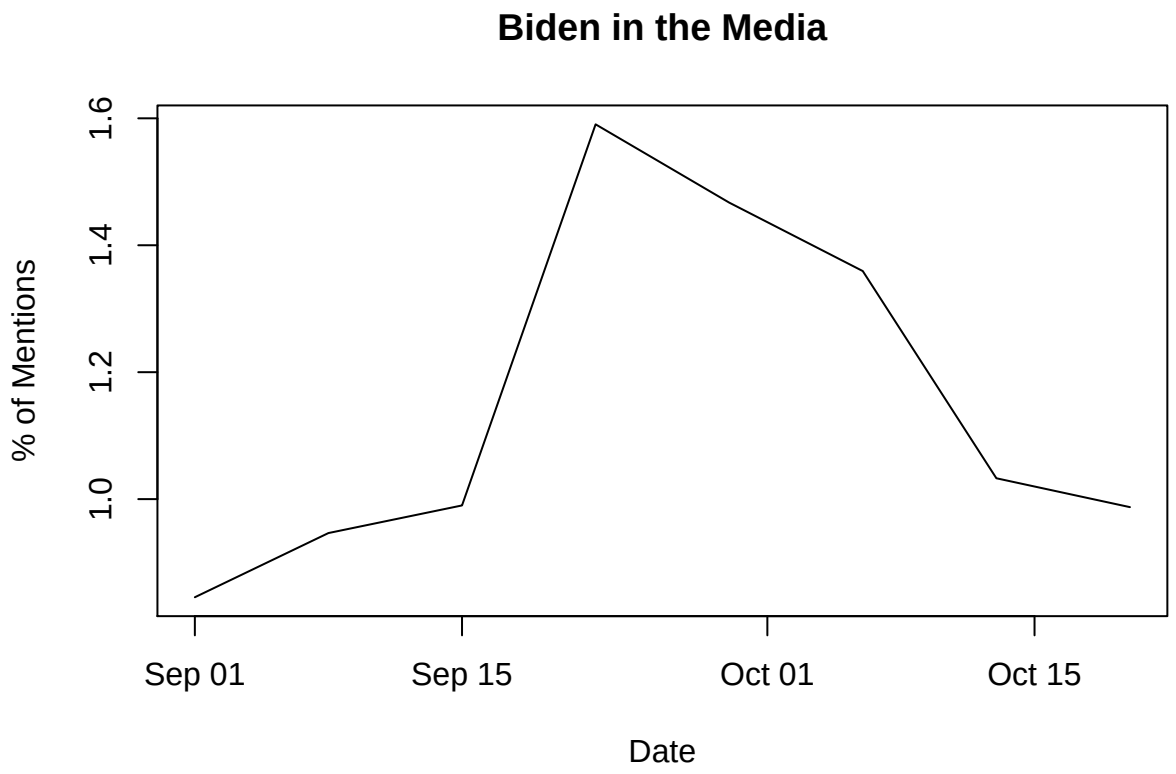
I started here by looking at the enitre set of data.

```
plot(all_biden$Date, all_biden$Clips + all_biden$Stories,
     type = "l", main = "Biden in the Media", xlab = "Date",
     ylab = "% of Mentions")
```



After looking at the data, it's clear that there is a sharp decrease in Joe Biden mentions starting somewhere in September With this in mind, I subsetted the data again, looking just at mentions from September on.

```
all_biden <- subset(all_biden, Date >= "2019-09-01")
plot(all_biden$Date, all_biden$Clips + all_biden$Stories,
     type = "l", main = "Biden in the Media", xlab = "Date",
     ylab = "% of Mentions")
```



Conclusions