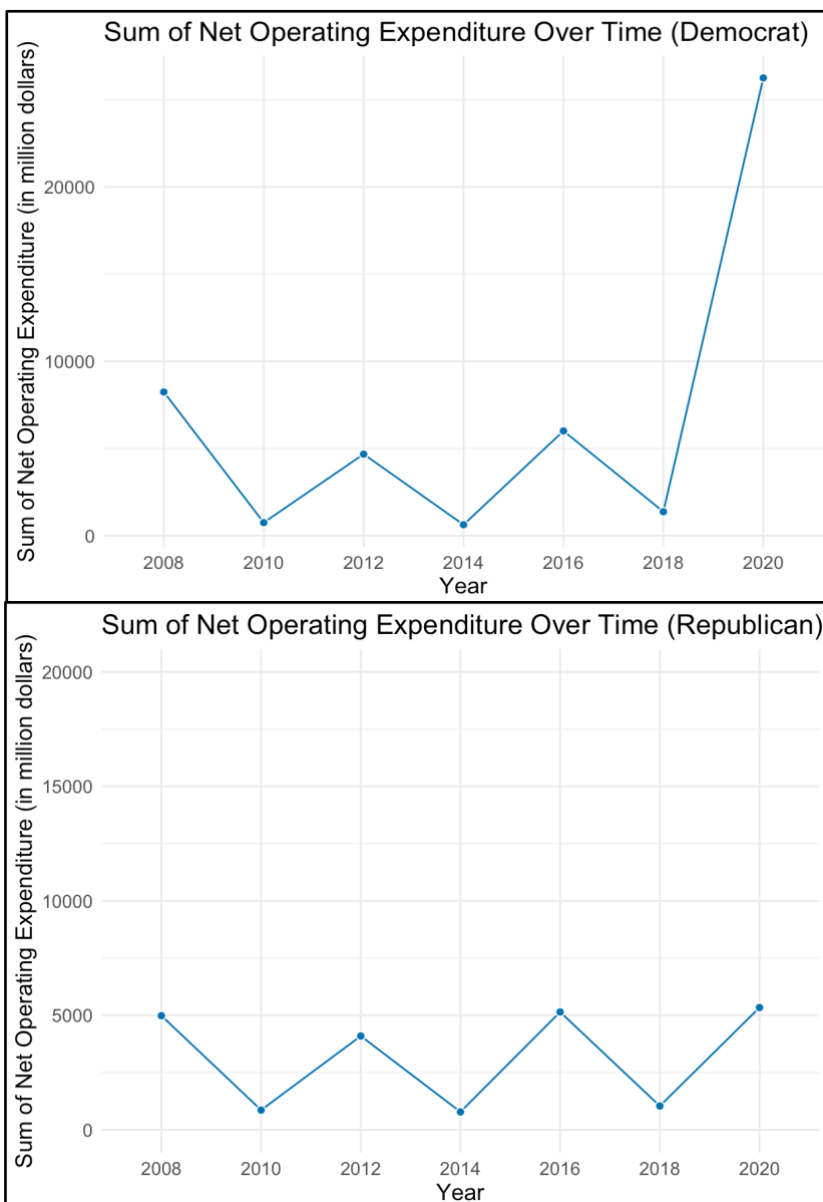


### Assignment 3: Misleading visualizations

#### PART 1

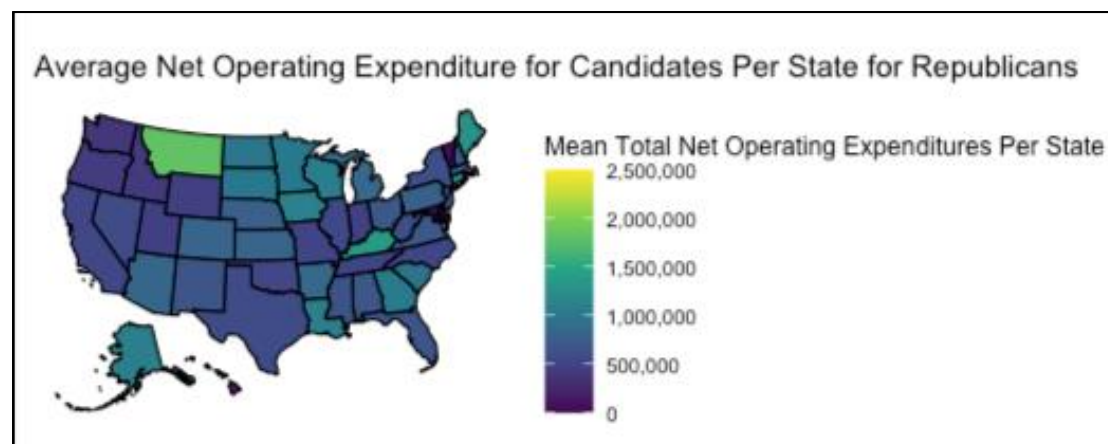
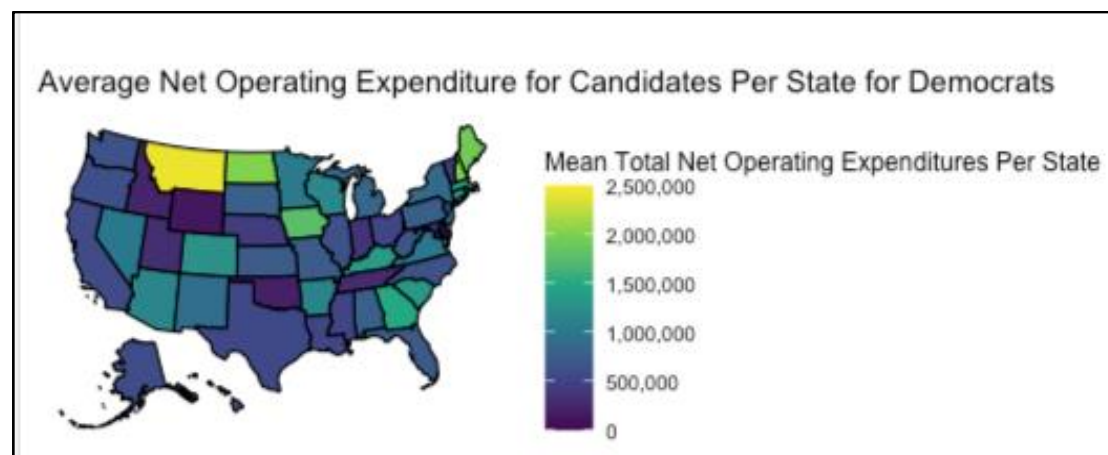
Democratic candidates across all parties spend more on campaigning efforts from 2008-2022 compared to their Republican counterparts. This can be seen by comparing the sum of Net Operating Expenditures between both parties. The two time series visualizations below, one for each political party, show the net operating expenditures over different election cycles for each party in millions of dollars. For each election cycle, there is a common trend that Democrats have spent significantly more than Republicans, with this trend increasing significantly in previous years. This supports the theory that Republican constituents are more supportive of their party since Republicans have to spend less on campaigning and spreading awareness during election season.



By comparing these two time series, we can make a few observations that will help develop insights on the data. A data point that stands out is that in 2008, Democratic candidates spent around \$7.5 billion in expenditures while Republican candidates spent 25% less on campaigning at \$5 billion. Another data point is that in 2016, Democrats in total spent around \$6 billion while Republicans spent right under \$5 billion, meaning that they spent around 16% less. Finally, in 2020, Democrats spent an astonishing \$12 billion while Republicans spent a relatively miniscule amount of around \$1 billion. All these points indicate that Republicans have been spending less for each cycle, which could indicate that Republicans need to spend less on convincing their constituents for support.

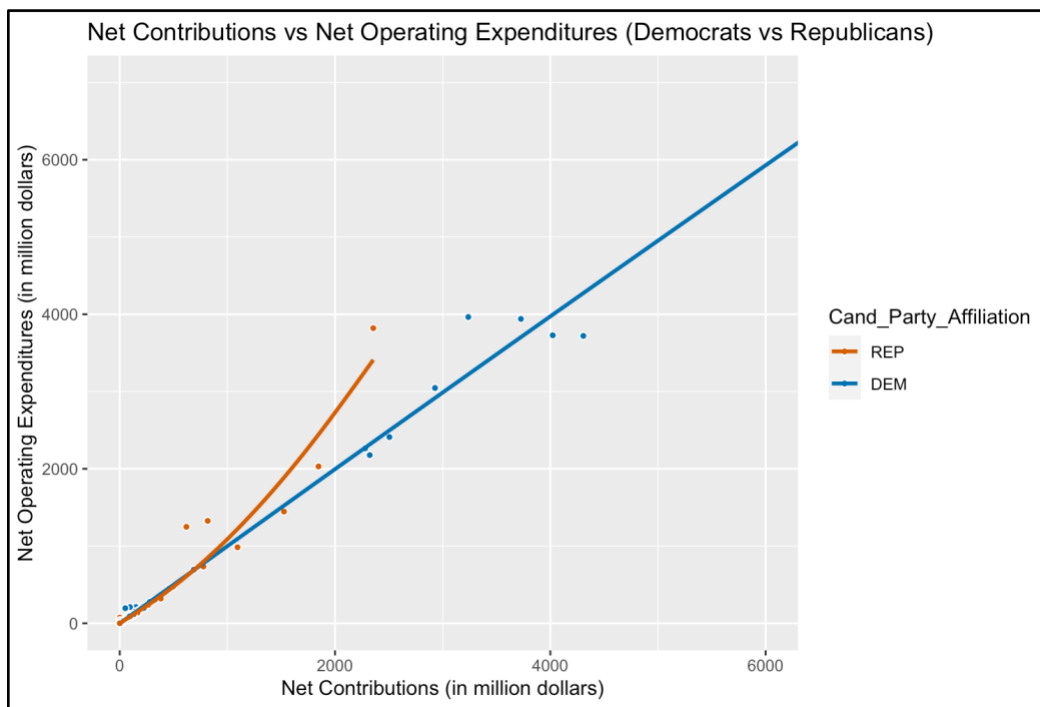
Nikhil Deshpande  
Viraj Vijaywargiya

The data can be looked at from a different perspective of a choropleth. The choropleth shows the net operating expenditures on average for each specific state and there are two separate choropleths for each political party. For Republicans, states such as Montana, Maine and Kentucky are the states that they spend the most in. These states seem to have high importance for Republicans to maintain incumbency in comparison to states such as Washington or California where they have spent little in comparison. For Democrats, states such as North Dakota, Iowa, Maine, and Montana have high importance for them to contest in. A reason for why there is a commonality between both political parties in where they spend most of their money could be because they are all swing states. These states have a split population, meaning that both parties need to heavily invest in these areas to represent them. States such as California will not require heavy investment from Democrats and states like Texas will not require heavy investment from Republicans. This is because both states already have a strong following for each respective party, so there is no need to campaign as much in these areas compared to more contentious ones. This perspective also shows us that Democrats are spending more in these split states compared to Republicans, which supports our theory.



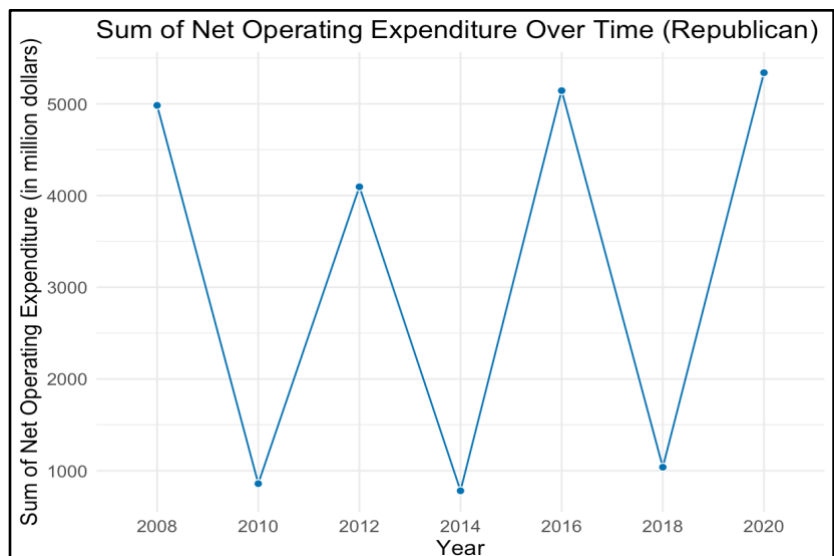
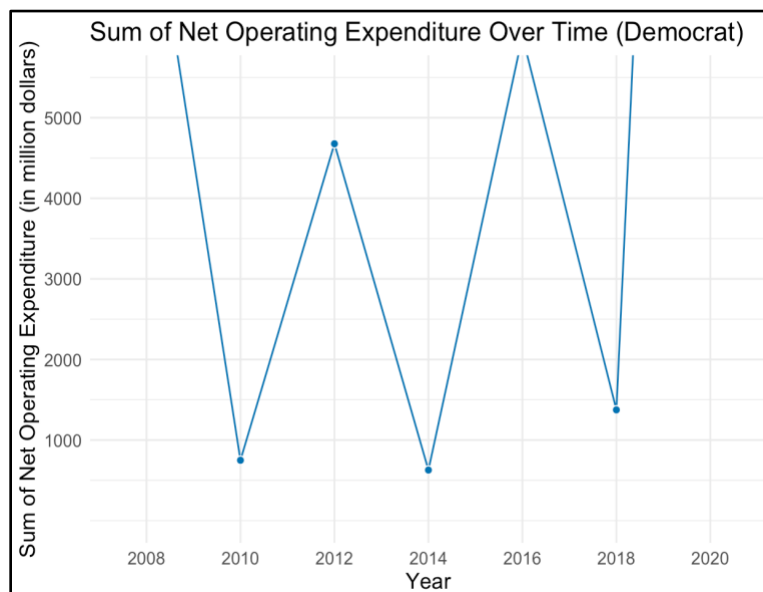
Nikhil Deshpande  
Viraj Vijaywargiya

A third way for us to interpret the data is by comparing each party's Net Contributions compared to their Operating Expenditures over different periods of time. We notice that in multiple time periods, Democrats have received significantly more net contributions and have correspondingly spent more on their campaigning efforts. The trend line for Democrats shows that they have been spending as much as their constituents have contributed to their campaigning efforts, meaning a  $x=y$  trend line. For Republicans, they seem to be spending more than what has been contributed to them by their constituents, meaning that loans are not taken into account. Republicans seem to depend less on campaign contributions from the public compared to Democrats. Overall, this visualization also leads us to believe that Democrats are spending more than Republicans, regardless of the contributions that are received by their constituents.



## PART 2

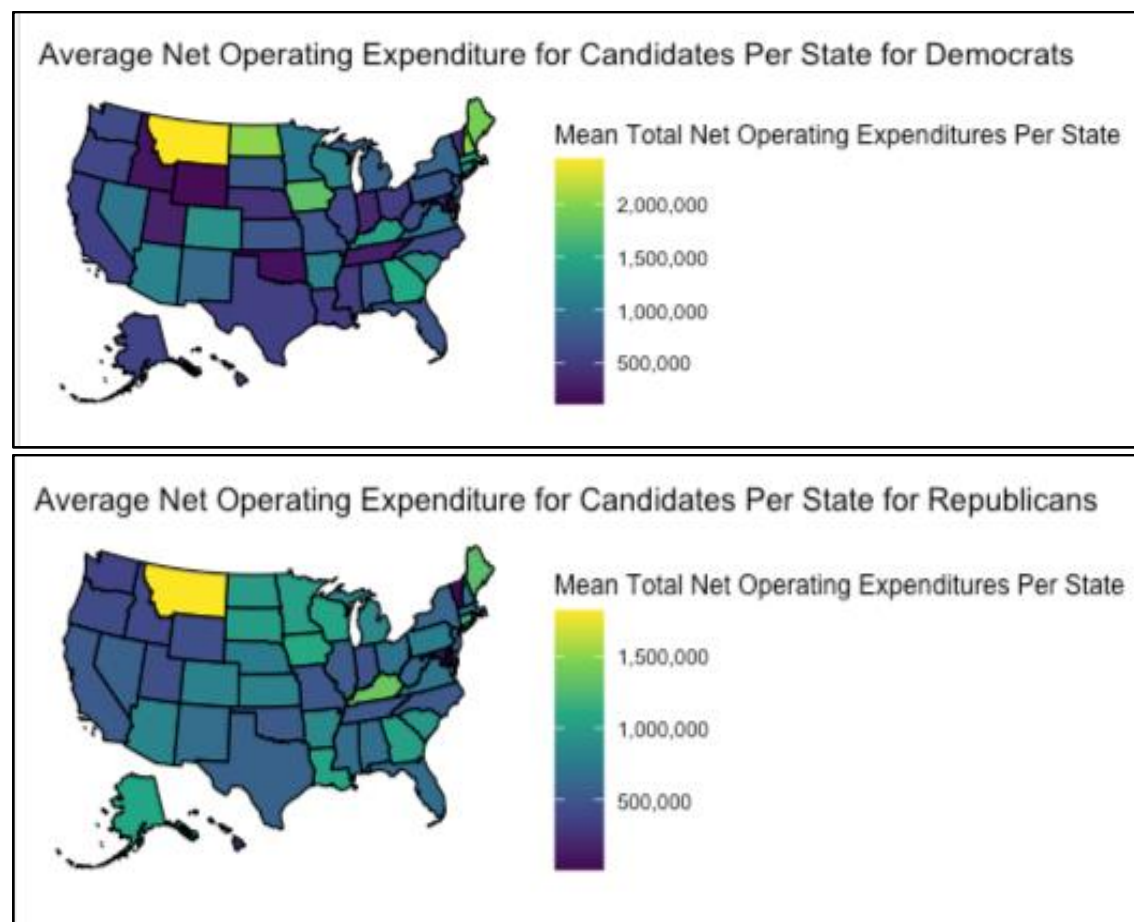
Democratic candidates across all parties spent almost the same amount for campaigning from 2008-2022 compared to their Republican counterparts. This can be seen by comparing the sum of Net Operating Expenditures between both parties. The two time series visualizations below, one for each political party, show the net operating expenditures over different election cycles for each party in millions of dollars. For each election cycle, both parties follow similar patterns of spending. A spike in spending for both parties is visible during the presidential campaign cycles, following a “W-shape” pattern. Both parties have similar maximum and minimum amounts, meaning that their spending strategies are very similar. For example, both parties spent nearly the same amount on campaigning during the years 2010, 2012, and 2018. This supports the theory that Republicans and Democrats have put in the same amount of campaigning efforts and investment to get support from their constituents.



Nikhil Deshpande

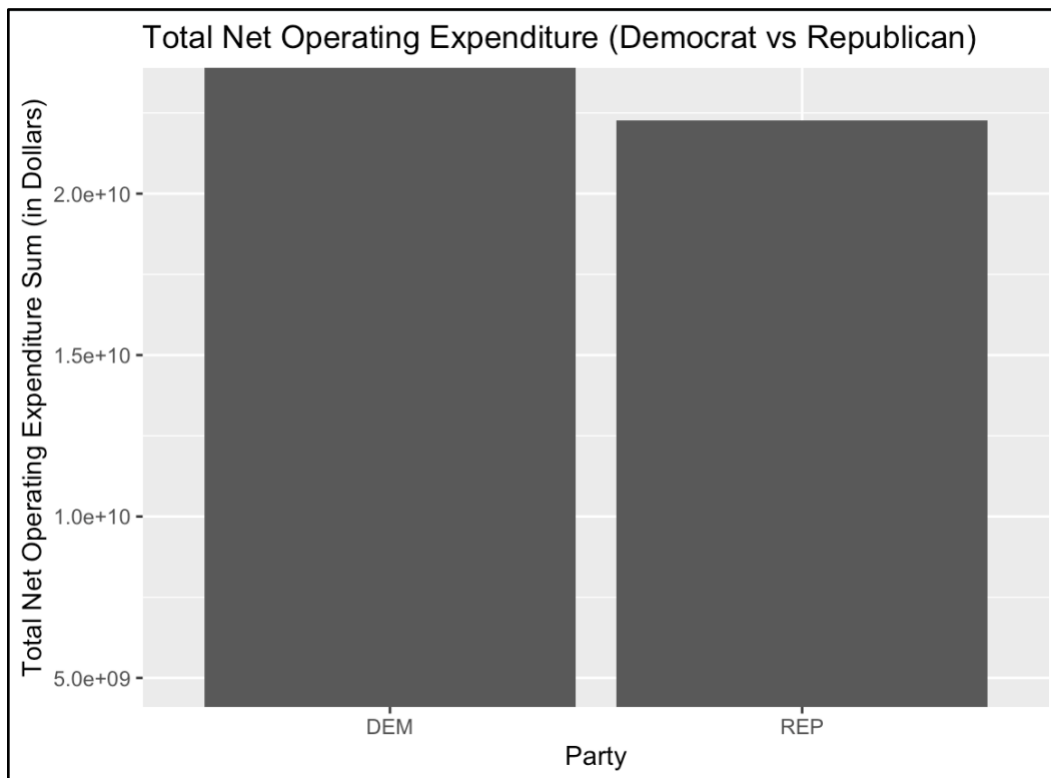
Viraj Vijaywargiya

The data can be looked at from a different perspective of a choropleth. The choropleth shows the net operating expenditures on average for each specific state and there are two separate choropleths for each political party. For Republicans, states such as Montana, Maine, and Kentucky are where they invest the most for campaigning. These states seem to have high importance for Republicans to maintain incumbency in comparison to states such as Washington or California where they have spent little in comparison. Democrats spend nearly the same amounts in the same states of Montana, Maine, and Kentucky, as seen by the same color grade. A reason for why there is a commonality between both political parties in where they spend most of their money could be because they are all swing states. These states have a split population, meaning that both parties need to heavily invest in these areas to represent them. States such as California will not require heavy investment from Democrats and states like Texas will not require heavy investment from Republicans. This is because both states already have a strong following for each respective party, so there is no need to campaign as much in these areas compared to more contentious ones. This perspective also shows us that Democrats are spending the same in these split states compared to Republicans, which supports our theory.



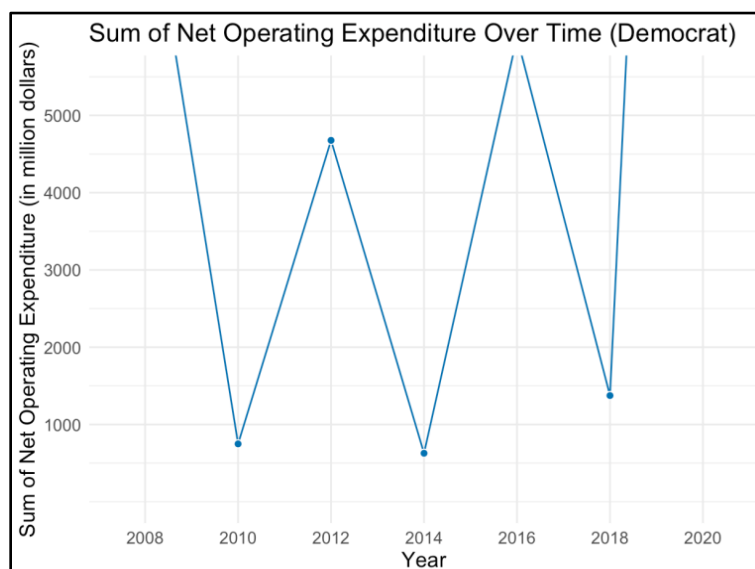
Nikhil Deshpande  
Viraj Vijaywargiya

A third way for us to interpret the data is by comparing each party's Net Operating Expenditures in aggregate through a bar graph. We notice overall, Democrats have received around the same, if not marginally more than Republicans for their campaigning efforts. Across all different offices, it seems that holistically the net expenditures are around the same between Republicans and Democrats. This means that both parties have similar spending strategies and both party's constituents require the same amount of campaigning investments to keep their support. Therefore, this visualization leads us to believe that Democrats and Republicans have almost identical spending behavior during election cycles.



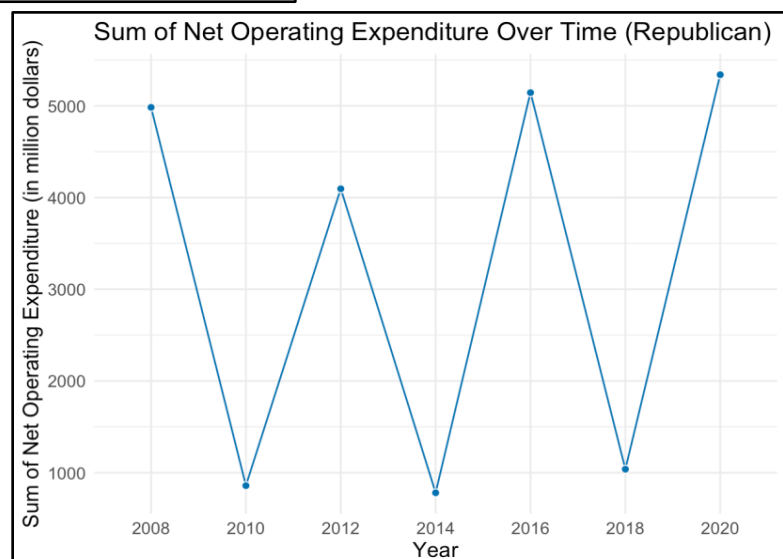
### PART 3 (Annotations are highlighted in yellow)

Democratic candidates across all parties spent almost the same amount for campaigning from 2008-2022 compared to their Republican counterparts. This can be seen by comparing the sum of Net Operating Expenditures between both parties. The two time series visualizations below, one for each political party, show the net operating expenditures over different election cycles for each party in millions of dollars. For each election cycle, both parties follow similar patterns of spending. A spike in spending for both parties is visible during the presidential campaign cycles, following a “W-shape” pattern. Both parties have similar maximum and minimum amounts, meaning that their spending strategies are very similar. For example, both parties spent nearly the same amount on campaigning during the years 2010, 2012, and 2018. This supports the theory that Republicans and Democrats have put in the same amount of campaigning efforts and investment to get support from their constituents.



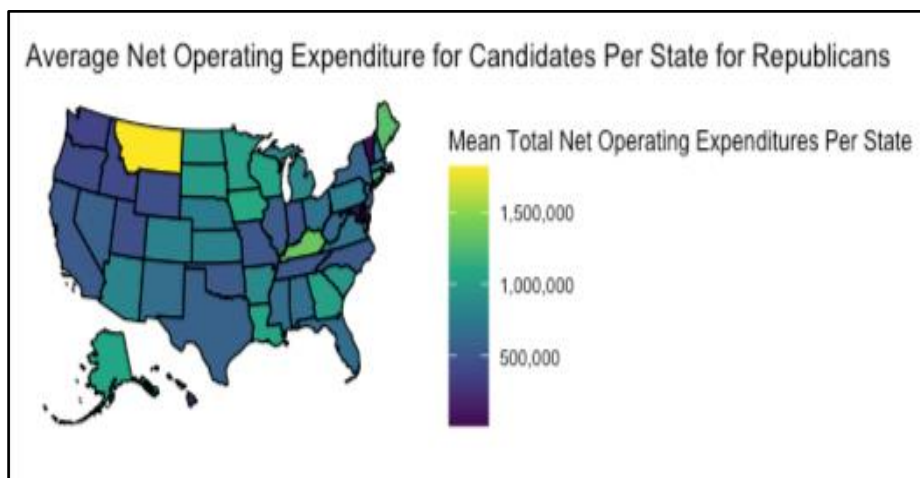
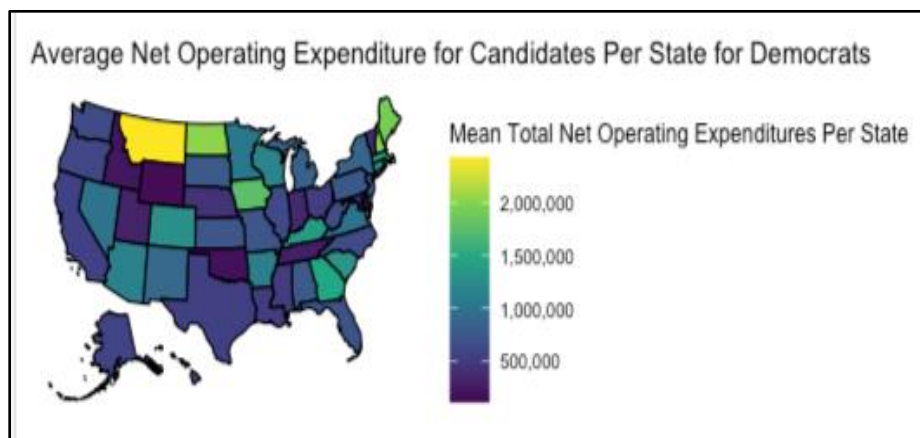
At first glance, both the graphs look normal but, if you look closely, the values for the sum of net operating expenditure for democrats in the years 2008, 2016, and 2020 exceed the frame. However, the values for the sum of net operating expenditure for republicans are all in the frame. This is because the y axis values for democrats were limited from 0 to 5500 which makes both the graphs look the same, however, they aren't.

This modification that resulted in a misleading visualization makes the story misleading too. From these two graphs, you can see that both democrats and republicans have almost the same trend. However, the values of democrats go upto 28000 which clearly means democrats are spending more than republicans.





The data can be looked at from a different perspective of a choropleth. The choropleth shows the net operating expenditures on average for each specific state and there are two separate choropleths for each political party. For Republicans, states such as Montana, Maine, and Kentucky are where they invest the most for campaigning. These states seem to have high importance for Republicans to maintain incumbency in comparison to states such as Washington or California where they have spent little in comparison. Democrats spend nearly the same amounts in the same states of Montana, Maine, and Kentucky, as seen by the same color grade. A reason for why there is a commonality between both political parties in where they spend most of their money could be because they are all swing states. These states have a split population, meaning that both parties need to heavily invest in these areas to represent them. States such as California will not require heavy investment from Democrats and states like Texas will not require heavy investment from Republicans. This is because both states already have a strong following for each respective party, so there is no need to campaign as much in these areas compared to more contentious ones. This perspective also shows us that Democrats are spending the same in these split states compared to Republicans, which supports our theory.

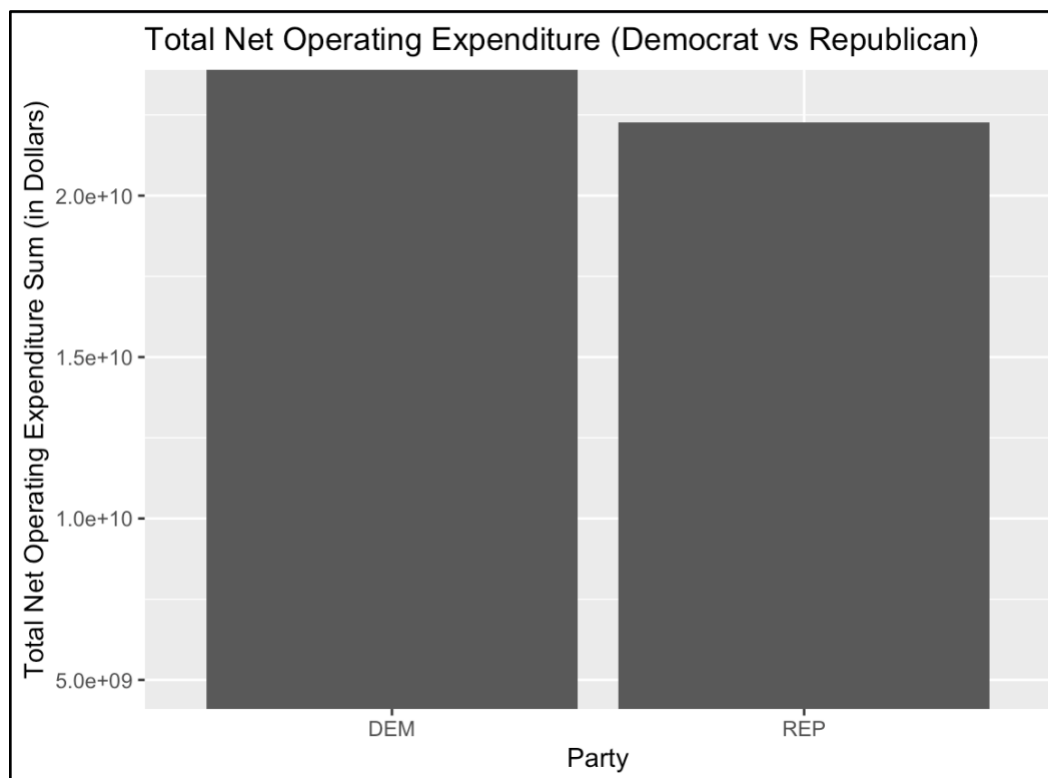


At first glance, the two choropleth graphs, for democrats and republicans, seem as if they almost have the same spending habits by states. However, looking closely at the color scale (legend), you can see that democrat's scale is set from 0 to 2500000, whereas republican's scale is set from 0 to 2000000. This makes the data misleading and does not show accurate densities of color per state that represent their mean total net operating expenditures.



Nikhil Deshpande  
Viraj Vijaywargiya

A third way for us to interpret the data is by comparing each party's Net Operating Expenditures in aggregate through a bar graph. We notice overall, Democrats have received around the same, if not marginally more than Republicans for their campaigning efforts. Across all different offices, it seems that holistically the net expenditures are around the same between Republicans and Democrats. This means that both parties have similar spending strategies and both party's constituents require the same amount of campaigning investments to keep their support. Therefore, this visualization leads us to believe that Democrats and Republicans have almost identical spending behavior during election cycles.



At first glance, you can interpret from the bar chart above that both the parties have almost the same spending habits as their total net operating expenditures aren't much different. However, looking at the chart closely, you can see that the chart does not start from the baseline of 0. Additionally, the y axis scale limits (5000000000 to 23000000000) are set so that the total net operating expenditure for democrats is cut and the real value is not shown. These modifications to the data creates a very misleading visualization, making us believe that democrats and republicans have almost the same total net operating expenditure. However, in reality, the value for democrats go upto 4.7e+10, which is around twice of that of the republicans.

Nikhil Deshpande

Viraj Vijaywargiya

### **Group Contribution**

#### **PART 1: Good Visualizations**

**Viraj Vijaywargiya** mainly created the following visualizations: Sum of Net Operating Expenditures Over Time (Democrat), Sum of Net Operating Expenditures Over Time (Republican), and Net Contributions Vs Net Operating Expenditures (Democrats vs Republicans).

**Nikhil Deshpande** mainly created the following visualizations: Average Net Expenditures for Candidates Per State for Republicans, and Average Net Expenditures for Candidates Per State for Democrats.

#### **PART 2: Misleading Visualizations**

**Viraj Vijaywargiya** mainly created the following visualizations: Sum of Net Operating Expenditures Over Time (Democrat), Sum of Net Operating Expenditures Over Time (Republican).

**Nikhil Deshpande** mainly created the following visualizations: Average Net Expenditures for Candidates Per State for Republicans, Average Net Expenditures for Candidates Per State for Democrats, and Total Net Operating Expenditure(Democrat vs Republicans).

Nikhil Deshpande  
Viraj Vijaywargiya

## **CODE:**

### **Good Visualizations**

```
library(ggplot2)
library(tidyverse)
library(usmap)
library(dplyr)
library(scales)

df <- read.csv("https://www.ics.uci.edu/~algol/teaching/informatics143w2022/fec_2008-2022.csv")

df <- df %<>%
  mutate(Year = substr(Coverage_End_Date, nchar(Coverage_End_Date)-4+1,
                        nchar(Coverage_End_Date)))
df$Coverage_Start_Date <- mdy(df$Coverage_Start_Date)
df$Coverage_Start_Date <- as.Date(df$Coverage_Start_Date)
df$Coverage_End_Date <- mdy(df$Coverage_End_Date)
df$Coverage_End_Date <- as.Date(df$Coverage_End_Date)
```

### **Visualization 1:**

```
DemTS <- df[which(df$Cand_Party_Affiliation=="DEM"),]
DemTS <- DemTS %>%
  group_by(Year) %>%
  summarize(sum = sum(Net_Operating_Expenditure))

DemTS <- DemTS[which(DemTS$sum > 10000000),]

ggplot(DemTS, aes(x = Year, y = sum/1000000, group = 1)) +
  geom_line(size = 0.5, color = "#0072B2") +
  geom_point(color = "white", fill = "#0072B2", shape = 21, size = 2) +
  ylab("Sum of Net Operating Expenditure (in million dollars)") +
  ggtitle("Sum of Net Operating Expenditure Over Time (Democrat)") +
  #scale_x_date(name = "year") +
  theme_minimal() +
  theme(text = element_text(size=13))
```

Nikhil Deshpande

Viraj Vijaywargiya

### Visualization 2:

```
RepTS <- df[which(df$Cand_Party_Affiliation=="REP"),]
RepTS <- RepTS %>%
  group_by(Year) %>%
  summarize(sum = sum(Net_Operating_Expenditure))

RepTS <- RepTS[which(RepTS$sum > 10000000),]

ggplot(RepTS, aes(x = Year, y = sum/1000000, group = 1)) +
  geom_line(size = 0.5, color = "#0072B2") +
  geom_point(color = "white", fill = "#0072B2", shape = 21, size = 2) +
  coord_cartesian(ylim = c(0, 20000)) +
  ylab("Sum of Net Operating Expenditure (in million dollars)") +
  ggtitle("Sum of Net Operating Expenditure Over Time (Republican)") +
  #scale_x_date(name = "year") +
  theme_minimal() +
  theme(text = element_text(size=13))
```

### Visualization 3:

```
dem_cand <- df[df$Cand_Party_Affiliation=="DEM", ]
rep_cand <- df[df$Cand_Party_Affiliation=="REP", ]

chloro1 <- dem_cand %>% group_by(Cand_Office_St) %>% summarize(mean =
mean(Net_Operating_Expenditure))
chloro1$state = chloro1$Cand_Office_St
plot_usmap(data=chloro1, values="mean",) + scale_fill_continuous(type="viridis", name =
"Mean Total Net Operating Expenditures Per State", label = scales::comma, limits = c(0,
3000000)) + theme(legend.position = "right") + ggtitle("Average Net Operating Expenditure for
Candidates Per State for Democrats")
```

### Visualization 4:

```
chloro2 <- rep_cand %>% group_by(Cand_Office_St) %>% summarize(mean =
mean(Net_Operating_Expenditure))
chloro2$state = chloro2$Cand_Office_St
plot_usmap(data=chloro2, values="mean",) + scale_fill_continuous(type="viridis", name =
"Mean Total Net Operating Expenditures Per State", label = scales::comma, limits = c(0,
3000000)) + theme(legend.position = "right") + ggtitle("Average Net Operating Expenditure for
Candidates Per State for Republicans")
```

Nikhil Deshpande  
Viraj Vijaywargiya

### Visualization 5:

```
df2 <- df[which((df$Cand_Party_Affiliation=="DEM" | df$Cand_Party_Affiliation=="REP")),]  
df2 <- df2[which(df2$Net_Contribution>0 & df2$Net_Operating_Expenditure>0),]  
  
ggplot(df2, aes(x=Net_Contribution/1000000, y=Net_Operating_Expenditure/1000000,  
fill=Cand_Party_Affiliation)) +  
  geom_point(pch = 21, color = "white") +  
  coord_cartesian(xlim=c(0, 6000), ylim = c(0, 7000)) +  
  geom_smooth(se = FALSE, method = "gam", formula = y ~ s(x, k = 3),  
aes(color=Cand_Party_Affiliation)) +  
  scale_fill_manual(values = c(REP = "#D55E00", DEM = "#0072B2")) +  
  scale_color_manual(values = c(REP = "#D55E00", DEM = "#0072B2")) +  
  labs(x = "Net Contributions (in million dollars)", y = "Net Operating Expenditures (in million  
dollars)") +  
  ggtitle("Net Contributions vs Net Operating Expenditures (Democrats vs Republicans)")
```

### Misleading Visualizations

```
library(ggplot2)  
library(tidyverse)  
library(usmap)  
library(dplyr)  
library(scales)
```

```
df <- read.csv("https://www.ics.uci.edu/~algol/teaching/informatics143w2022/fec_2008-  
2022.csv")
```

```
df <- df %<>%  
  mutate(Year = substr(Coverage_End_Date, nchar(Coverage_End_Date)-4+1,  
                        nchar(Coverage_End_Date)))  
df$Coverage_Start_Date <- mdy(df$Coverage_Start_Date)  
df$Coverage_Start_Date <- as.Date(df$Coverage_Start_Date)  
df$Coverage_End_Date <- mdy(df$Coverage_End_Date)  
df$Coverage_End_Date <- as.Date(df$Coverage_End_Date)
```

### Visualization 1:

```
DemTS <- df[which(df$Cand_Party_Affiliation=="DEM"),]  
DemTS <- DemTS %>%  
  group_by(Year) %>%  
  summarize(sum = sum(Net_Operating_Expenditure))
```

Nikhil Deshpande  
Viraj Vijaywargiya

```
DemTS <- DemTS[which(DemTS$sum > 10000000),]
```

```
view(DemTS)
```

```
ggplot(DemTS, aes(x = Year, y = sum/1000000, group = 1)) +  
  geom_line(size = 0.5, color = "#0072B2") +  
  geom_point(color = "white", fill = "#0072B2", shape = 21, size = 2) +  
  coord_cartesian(ylim = c(0, 5500)) +  
  scale_y_continuous(breaks = c(1000, 2000, 3000, 4000, 5000)) +  
  ylab("Sum of Net Operating Expenditure (in million dollars)") +  
  ggtitle("Sum of Net Operating Expenditure Over Time (Democrat)") +  
  theme_minimal() +  
  theme(text = element_text(size=13))
```

### Visualization 2:

```
RepTS <- df[which(df$Cand_Party_Affiliation=="REP"),]  
RepTS <- RepTS %>%  
  group_by(Year) %>%  
  summarize(sum = sum(Net_Operating_Expenditure))
```

```
RepTS <- RepTS[which(RepTS$sum > 10000000),]
```

```
ggplot(RepTS, aes(x = Year, y = sum/1000000, group = 1)) +  
  geom_line(size = 0.5, color = "#0072B2") +  
  geom_point(color = "white", fill = "#0072B2", shape = 21, size = 2) +  
  ylab("Sum of Net Operating Expenditure (in million dollars)") +  
  ggtitle("Sum of Net Operating Expenditure Over Time (Republican)") +  
  theme_minimal() +  
  theme(text = element_text(size=13))
```

### Visualization 3:

```
dem_cand <- df[df$Cand_Party_Affiliation=="DEM", ]  
rep_cand <- df[df$Cand_Party_Affiliation=="REP", ]
```

```
chloro1 <- dem_cand %>% group_by(Cand_Office_St) %>% summarize(mean =  
  mean(Net_Operating_Expenditure))  
chloro1$state = chloro1$Cand_Office_St  
plot_usmap(data=chloro1, values="mean",) + scale_fill_continuous(type="viridis", name =  
  "Mean Total Net Operating Expenditures Per State", label = scales::comma) +  
  theme(legend.position = "right") + ggtitle("Average Net Operating Expenditure for Candidates  
  Per State for Democrats")
```

Nikhil Deshpande  
Viraj Vijaywargiya

#### Visualization 4:

```
chloro2 <- rep_cand %>% group_by(Cand_Office_St) %>% summarize(mean =  
mean(Net_Operating_Expenditure))  
chloro2$state = chloro2$Cand_Office_St  
plot_usmap(data=chloro2, values="mean",) + scale_fill_continuous(type="viridis", name =  
"Mean Total Net Operating Expenditures Per State", label = scales::comma) +  
theme(legend.position = "right") + ggtitle("Average Net Operating Expenditure for Candidates  
Per State for Republicans")
```

#### Visualization 5:

```
misDF <- read.csv("https://www.ics.uci.edu/~algol/teaching/informatics143w2022/fec_2008-  
2022.csv")
```

```
mis1 <- misDF[which((df$Cand_Party_Affiliation=="DEM" |  
df$Cand_Party_Affiliation=="REP")),]
```

```
mis1 <- mis1 %>%  
  group_by(Cand_Party_Affiliation) %>%  
  summarize(sum = sum(Net_Operating_Expenditure))
```

```
ggplot(data=mis1, aes(x=Cand_Party_Affiliation, y=sum)) +  
  geom_bar(stat="identity") +  
  scale_y_continuous(limits = c(5000000000, 23000000000), oob = rescale_none) +  
  labs(x = "Party", y="Total Net Operating Expenditure Sum (in Dollars)") +  
  ggtitle("Total Net Operating Expenditure (Democrat vs Republican)")
```