# SDS192 Mini-Project 2

## Monday, November 1st, 2021

## Administrative info:

- Section number: 01
- Group number: 41
- Group leader name: Karen Galvan
- Group leader academic email:kvgalvan@smith.edu
- Partner name: Isa West
- Partner academic email:iwest@smith.edu

## Peer evaluation Google Form

Fill out the peer evaluation Google Form to receive credit.

## Summary

Write no more than 100 words here summarizing how you are "following the money":

We were interested in following the money trail associated with one of the most important factors in determining the outcome of the 2016 presidential election: the election results in the swing states. We researched the ten most contested states in the election: Colorado, Florida, Iowa, Michigan, North Carolina, New Hampshire, Nevada, Ohio, Virginia, and Wisconsin[1]. Next, we examined both the average amount donated to the candidates (Clinton and Trump) in each state, and who the winner was. We then compared the total amounts donated, and the types of entities that donated, to each candidate in the states they won.

## Section 1 - Mean Contribution to Each Candidate by the Swing States

Write no more than 250 words describing and interpreting your first visualization here.

Here we visualized the mean contributions given to each candidate by swing state, with an emphasis on the election result. We expected to see that the candidate that received the largest average donations in each state often also won in that state. Our reasoning was that high average contribution amounts could indicate a large amount of support for that candidate among the population. While we observed this trend in five of the states, the other five contradicted this theory. In Colorado, North Carolina, New Hampshire, Virginia, and Wisconsin, the candidate that received the largest average contributions was not who won. This was most notable in Wisconsin where the average contribution was so insignificant for Trump in comparison to Clinton that it is barely visible on the graph, and yet ultimately he won the state. One reason for this could be that Trump received a lot of small individual donations, driving down the mean, despite resulting in a

---

[1]Mahtesian, Charlie. "What Are the Swing States in 2016?" POLITICO, 15 June 2016, https://www.politico.com/blogs/swing-states-2016-election/2016/06/what-are-the-swing-states-in-2016-list-224327

large total. After doing research, we learned that while the FEC tracks the sum of all contributions, they only record donations of more than $200 as distinct contributions. In the 2016 election, Trump raised money in a much shorter time frame than the other candidates, which reduced the likelihood that donors would make multiple donations, eventually reaching the $200 threshold[2]. These large amounts of small contributions that bring down, or aren't even represented in the mean, could explain the results we obtained for the other 4 states that contradicted our initial hypothesis.

```r
# Put code to create first data visualization here:

# combines contributions and candidates

contributions <- read_all_contributions()

joined_contributions <- candidates %>%
    inner_join(contributions, by = "cand_id")

filter_contributions <- joined_contributions %>%
    filter((cand_id == "P00003392" | cand_id == "P80001571") &
        (state == "FL" | state == "OH" | state == "CO" | state ==
            "NH" | state == "NC" | state == "IA" | state == "MI" |
            state == "NV" | state == "VA" | state == "WI"))

results_contributions <- filter_contributions %>%
    left_join(results_president, by = "cand_id") %>%
    select(cand_id, cand_name, cand_pty_affiliation, state.x,
        transaction_amt, entity_tp)

fixed_hillary <- results_contributions
fixed_hillary[results_contributions == "CLINTON, HILLARY RODHAM / TIMOTHY MICHAEL KAINE"] <- "Clinton"

fixed_trump <- fixed_hillary
fixed_trump[fixed_hillary == "TRUMP, DONALD J. / MICHAEL R. PENCE "] <- "Trump"

summary_results <- fixed_trump %>%
    group_by(cand_id, cand_name, cand_pty_affiliation, state.x) %>%
    summarize(mean = mean(transaction_amt, na.rm = TRUE))

final_summary <- summary_results
final_summary$results <- c("won", "lost", "lost", "lost", "lost",
    "won", "won", "lost", "won", "lost", "lost", "won", "won",
    "won", "won", "lost", "lost", "won", "lost", "won")


ggplot(data = final_summary, mapping = aes(x = cand_name, y = mean,
    color = results, fill = cand_pty_affiliation)) + geom_col(size = 1) +
    facet_wrap(~state.x, scales = "free") + scale_color_manual(values = c("grey86",
    "springgreen")) + scale_fill_manual(values = c("blue", "red")) +
    xlab("Candidate") + ylab("Mean Contribution (in US Dollars)") +
    labs(title = "Mean Contribution to Each Candidate by the Swing States",
        fill = "Candidate Party")
```
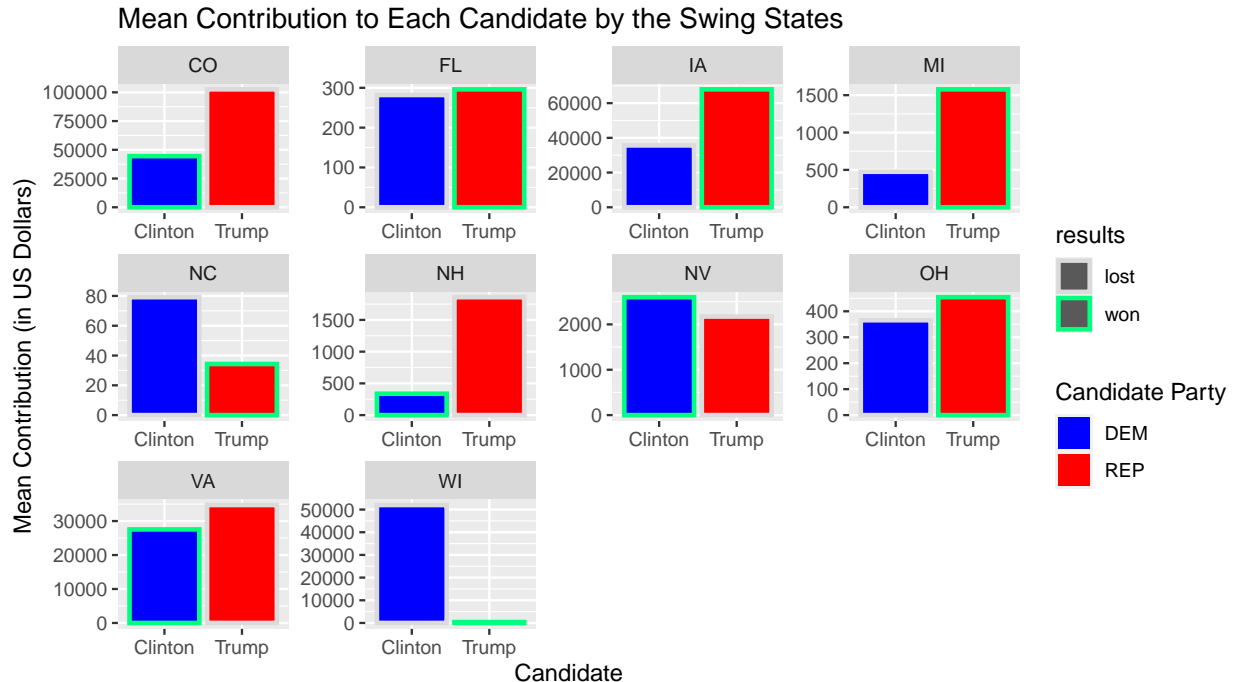
---

[2]Kayleigh McEnanystated on October, et al. "Politifact - Trump Raised More Dollars from Small Donations." @Politifact, https://www.politifact.com/factchecks/2017/nov/13/kayleigh-mcenany/trump-raised-more-dollars-small-donations/

## Mean Contribution to Each Candidate by the Swing States



# Section 2 Total Transactions by Entity in the Swing States Each Candidate Won

Write no more than 250 words describing and interpreting your second visualization here.

Here we visualized the amount contributed by entity type to the respective candidates in the states they won. The visualizations show that for both Trump and Clinton in the states they won, the entity type that consistently contributed the most were organizations. This might be, because 'organizations' is a broad category encompassing a wide variety of different groups, while other entity types have relatively narrower definitions. We could also have obtained these results, because organizations are able to donate much larger amounts to candidates than individuals are. It is also important to note that the amount of total individual contributions was likely much bigger, but was skewed by the FEC's $200 threshold discussed earlier[3]. In terms of total contributions, there was more variation between the states themselves, than between the two candidates, with both receiving relatively larger or smaller amounts depending on the state. Wisconsin however, contributed significantly less than any of the other swing states. This could be, because Wisconsin tends to lean Democratic, so Trump may have had less support there than in swing states that lean more Republican. When looking at the visualizations, we were surprised to see that Political Action Committees appeared only on the Virginia graph. After looking back at the original dataset, we discovered that there just didn't happen to be PAC contributions in the swing states. This also explains the absence of 'Committees' and 'Candidate Committees' on many of the graphs. Note: contributions that lacked an entity type were categorized under "NA."

```
# Put code to create second non-redundant data
# visualization here:

entity_summary_results <- fixed_trump %>%
    group_by(cand_id, cand_name, cand_pty_affiliation, state.x,
```

[3]Kayleigh McEnanystated on October, et al. "Politifact - Trump Raised More Dollars from Small Donations." @Politifact, https://www.politifact.com/factchecks/2017/nov/13/kayleigh-mcenany/trump-raised-more-dollars-small-donations/

```
        entity_tp) %>%
    filter(cand_name == "Trump" & (state.x == "FL" | state.x ==
        "OH" | state.x == "NC" | state.x == "IA" | state.x ==
        "MI" | state.x == "WI")) %>%
    summarize(entity_transactions = sum(transaction_amt, na.rm = TRUE))


ggplot(entity_summary_results, aes(x = entity_tp, y = entity_transactions,
    fill = cand_pty_affiliation)) + geom_col(size = 1) + facet_wrap(~state.x,
    scales = "free") + scale_fill_manual(values = c("red")) +
    xlab("Entity Type") + ylab("Total Transaction Amount (in US Dollars)") +
    labs(title = "Total Transactions by Entity in the Swing States Trump Won",
        fill = "Candidate Party")
```



Total Transactions by Entity in the Swing States Trump Won
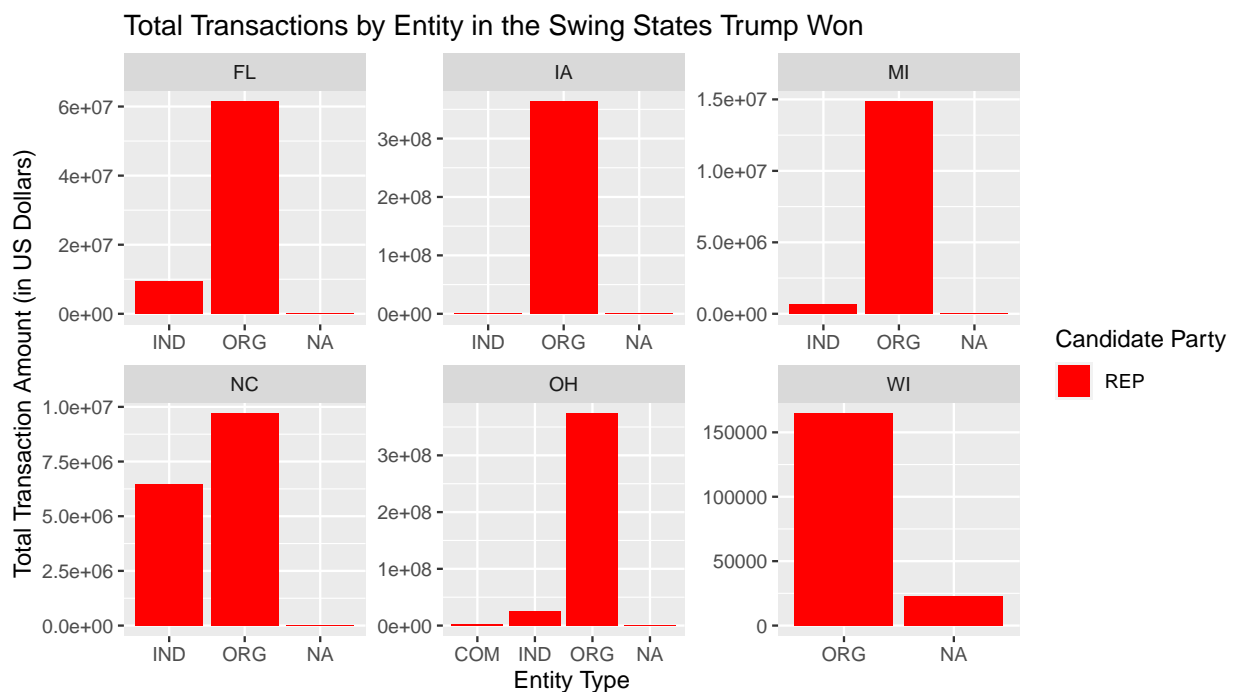
```
entity_summary_results2 <- fixed_trump %>%
    group_by(cand_id, cand_name, cand_pty_affiliation, state.x,
        entity_tp) %>%
    filter(cand_name == "Clinton" & (state.x == "CO" | state.x ==
        "NH" | state.x == "NV" | state.x == "VA")) %>%
    summarize(entity_transactions = sum(transaction_amt, na.rm = TRUE))


ggplot(entity_summary_results2, aes(x = entity_tp, y = entity_transactions,
    fill = cand_pty_affiliation)) + geom_col(size = 1) + facet_wrap(~state.x,
    scales = "free") + scale_fill_manual(values = c("blue")) +
    xlab("Entity Type") + ylab("Total Transaction Amount (in US Dollars)") +
    labs(title = "Total Transactions by Entity in the Swing States Clinton Won",
        fill = "Candidate Party")
```

4

Total Transactions by Entity in the Swing States Clinton Won
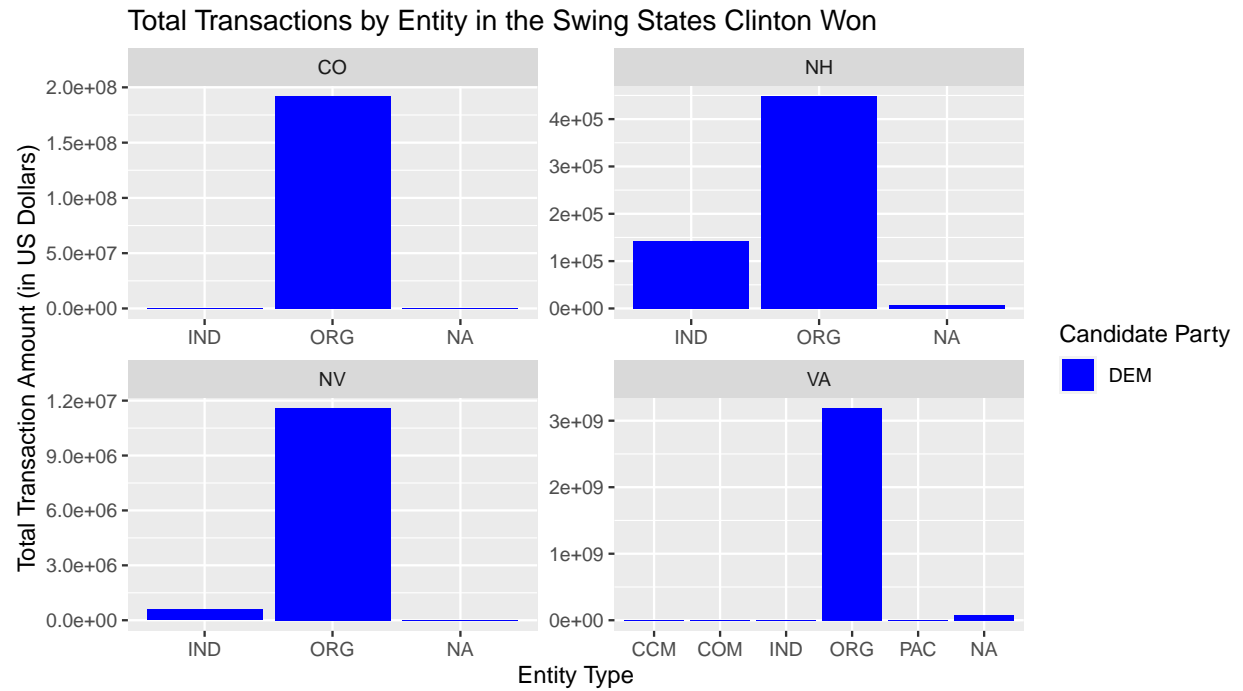


## Additional Analysis

(Strictly optional) Add any additional material here. [4]

---

[4]Here is an example footnote to be used as a citation.