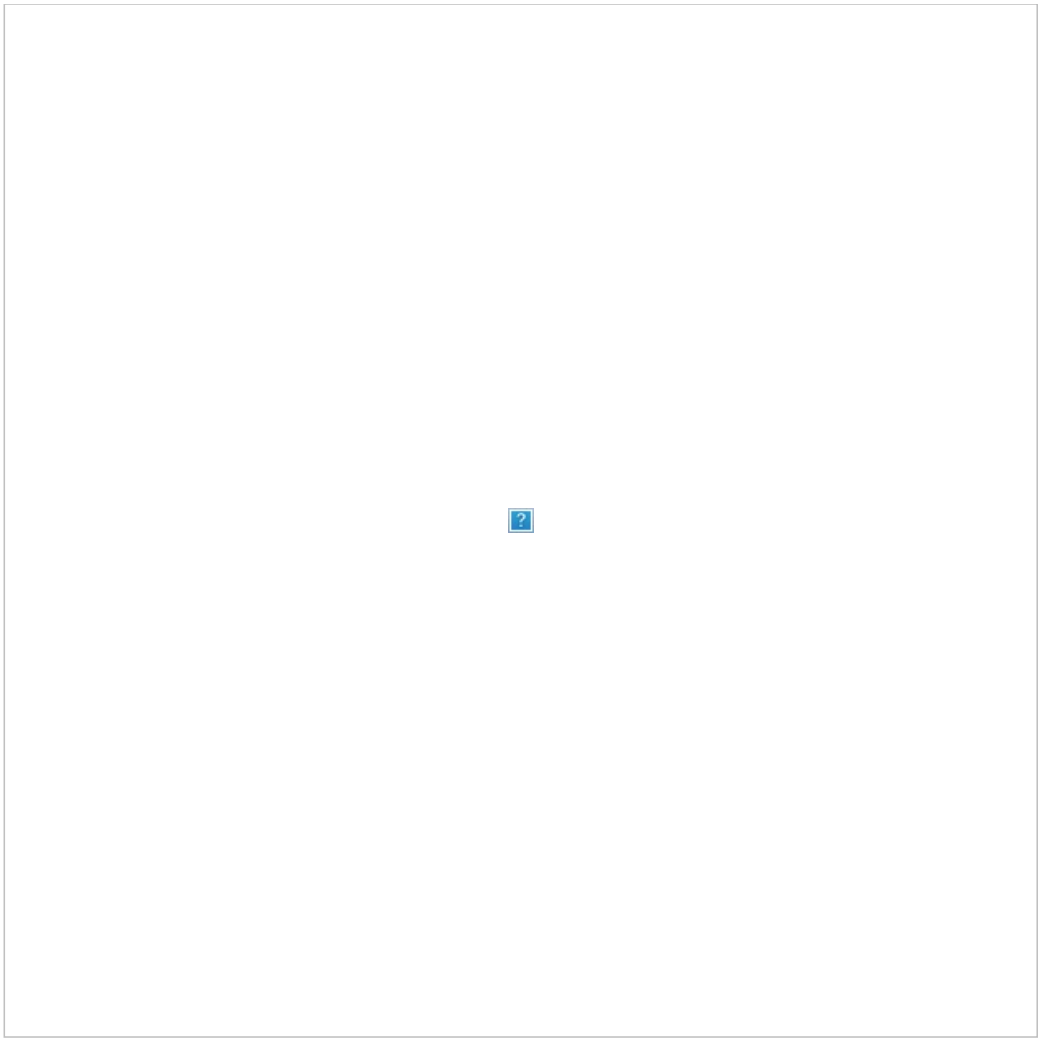
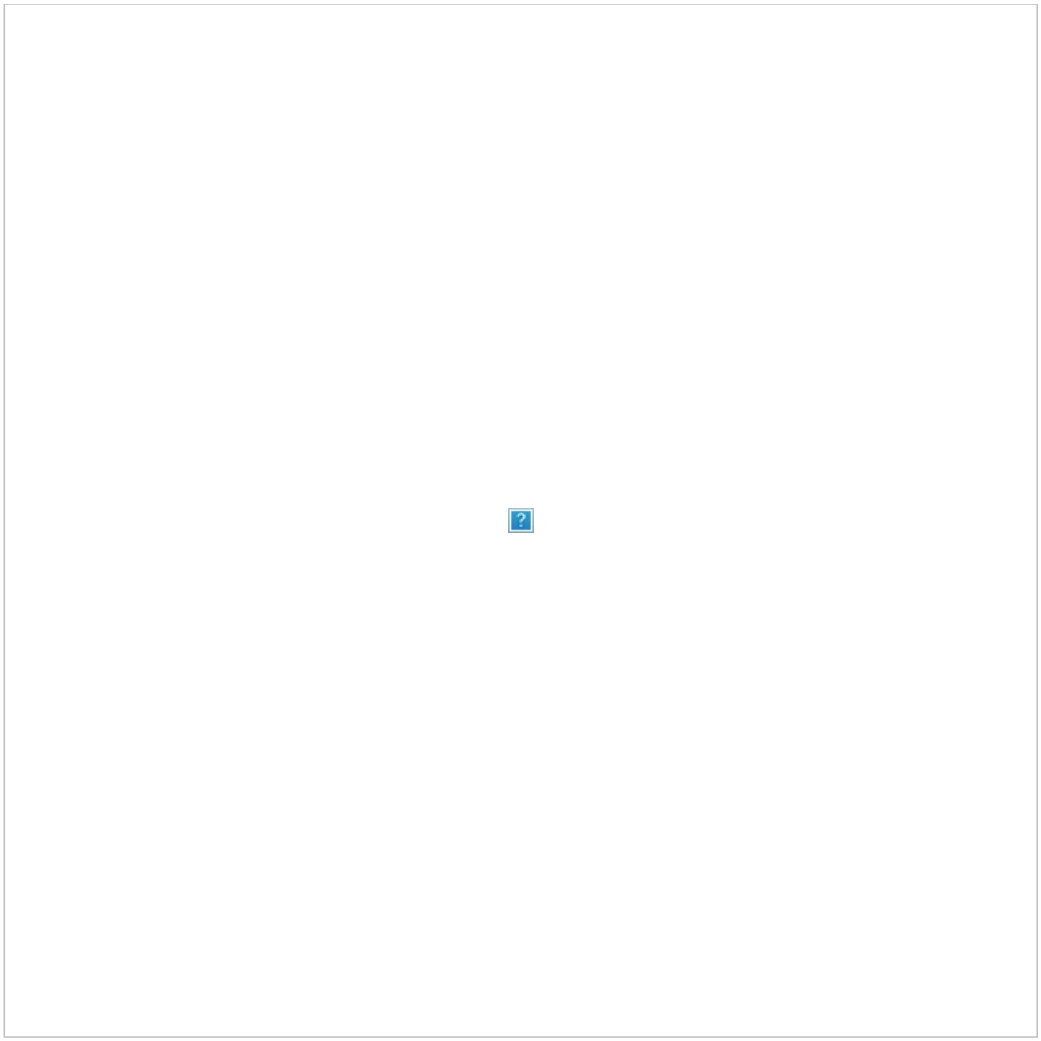
Final_Project_1st_Half

```
library(rio)
library(igraph)
library(dplyr)
library(readr)
library(tibble)
#https://stackoverflow.com/questions/52078017/create-similarity-matrix
library(data.table)


#Loading all the CA House and Senate data

#Each csv file in Edge_Lists represents 1 network. The first two columns are the sender/receiver identified with their "EID"
        (the identifier provided by NIMP). The third column is used to identify what threshold the edge was

CA_2011_2012_Senate_edges <- read_csv("Edgelist/CA-2011-2012-Senate.csv")
CA_2011_2012_Senate <- read_csv("Metadata/CA-2011-2012-Senate.csv")

CA_2013_2014_Senate_edges <- read_csv("Edgelist/CA-2013-2014-Senate.csv")
CA_2013_2014_Senate <- read_csv("Metadata/CA-2013-2014-Senate.csv")

CA_2015_2016_Senate_edges <- read_csv("Edgelist/CA-2015-2016-Senate.csv")
CA_2015_2016_Senate <- read_csv("Metadata/CA-2015-2016-Senate.csv")

CA_2011_2012_House <- read_csv("Metadata/CA-2011-2012-House.csv")
CA_2011_2012_House_edges <- read_csv("Edgelist/CA-2011-2012-House.csv")

CA_2013_2014_House <- read_csv("Metadata/CA-2013-2014-House.csv")
CA_2013_2014_House_edges <- read_csv("Edgelist/CA-2013-2014-House.csv")

CA_2015_2016_House <- read_csv("Metadata/CA-2015-2016-House.csv")
CA_2015_2016_House_edges <- read_csv("Edgelist/CA-2015-2016-House.csv")


#function to clean the data for visualizations

clean_data <- function(nodes, edges) {
  house_data_1 <- nodes %>%
    arrange(desc("Total")) %>%
    slice(1:20) %>%
    mutate(node_1 = EID) %>%
    select(-EID, -DemCol, -RepCol) %>%
    inner_join(edges, by = c("node_1")) %>%
    rename(Total_1 = Total) %>%
    mutate(name_1 = ContributorName) %>%
    mutate(group_1 = CatCodeGroup) %>%
    mutate(PerDem_1 = PerDem)

  house_data_2 <- nodes %>%
    arrange(desc("Total")) %>%
    slice(1:20) %>%
    mutate(node_2 = EID) %>%
    select(-EID, -DemCol, -RepCol) %>%
    inner_join(edges, by = c("node_2")) %>%
    rename(Total_2 = Total) %>%
    mutate(name_2 = ContributorName) %>%
    mutate(group_2 = CatCodeGroup) %>%
    mutate(PerDem_2 = PerDem) %>%
    select(node_2, Total_2, name_2, group_2, PerDem_2)

  full_house <- house_data_1 %>%
    inner_join(house_data_2, by = c("node_2")) %>%
    mutate(Total = Total_1 + Total_2) %>%
    arrange(desc(Total))

   return(full_house)
}

#function to create network of top 50 total donations between any donor tie based on yearly data
create_network <- function(df) {
  df <- df %>%
    distinct(name_1, name_2, edge, Total, group_1, group_2, PerDem_1, PerDem_2)

  house_net <- graph_from_data_frame(df, directed = T) #%>%

    #Democrat
  E(house_net)$color[which(E(house_net)$PerDem_1 >= 50 & E(house_net)$PerDem_2 >= 50)] <- rgb(0, 0, 1, .5)
    #Republican
  E(house_net)$color[which(E(house_net)$PerDem_1 < 50 & E(house_net)$PerDem_2 < 50)] <- rgb(1, 0, 0, .5)
```

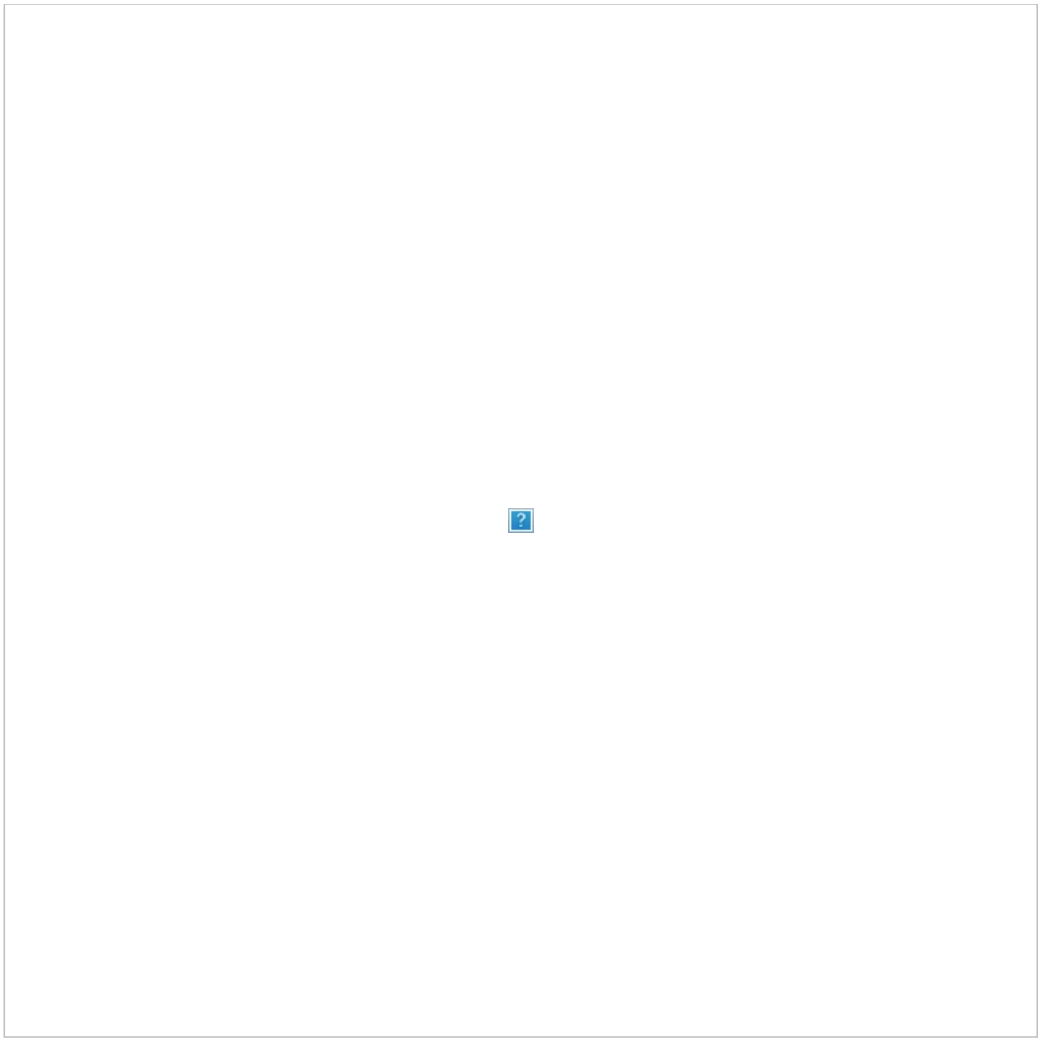
```
  return(house_net)
}

#CA House (2011-2016) data and network

House_11_12 <- clean_data(CA_2011_2012_House, CA_2011_2012_House_edges)
House_11_12_network <- create_network(House_11_12)

House_13_14 <- clean_data(CA_2013_2014_House, CA_2013_2014_House_edges)
House_13_14_network <- create_network(House_13_14)

House_15_16 <- clean_data(CA_2015_2016_House, CA_2015_2016_House_edges)
House_15_16_network <- create_network(House_15_16)


#CA Senate (2011-2018) data and network

Senate_11_12 <- clean_data(CA_2011_2012_Senate, CA_2011_2012_Senate_edges)
Senate_11_12_network <- create_network(Senate_11_12)

Senate_13_14 <- clean_data(CA_2013_2014_Senate, CA_2013_2014_Senate_edges)
Senate_13_14_network <- create_network(Senate_13_14)

Senate_15_16 <- clean_data(CA_2015_2016_Senate, CA_2015_2016_Senate_edges)
Senate_15_16_network <- create_network(Senate_15_16)


#Community Detection Algorithm for House 2011-2016

#We want to see how the top echelon of political donors are grouped to detect any patterns

#Optimal Community Detection Algorithm to maximize modularity score increasing cohesiveness in the groups

#Network modularity to determine level of polarization per year especially among the more avid donors to see if they exhibit
        a tendency to only interact within their own party group

#Network Modularity = 0.19
House_11_12_opt <- cluster_optimal(House_11_12_network)
#House_11_12_opt_plot <- plot(House_11_12_opt, House_11_12_network, vertex.color = V(House_11_12_network)$color, vertex.size
        = 10) + title(main = "Network Modularity: ", sub = modularity(House_11_12_opt))
House_11_12_opt_plot <- plot(House_11_12_opt, House_11_12_network, vertex.label = NA, vertex.color = "black", edge.color =
        E(House_11_12_network)$color, vertex.size = 15) + title(main = "CA House 11-12 Network Modularity: ", sub =
        modularity(House_11_12_opt))
```



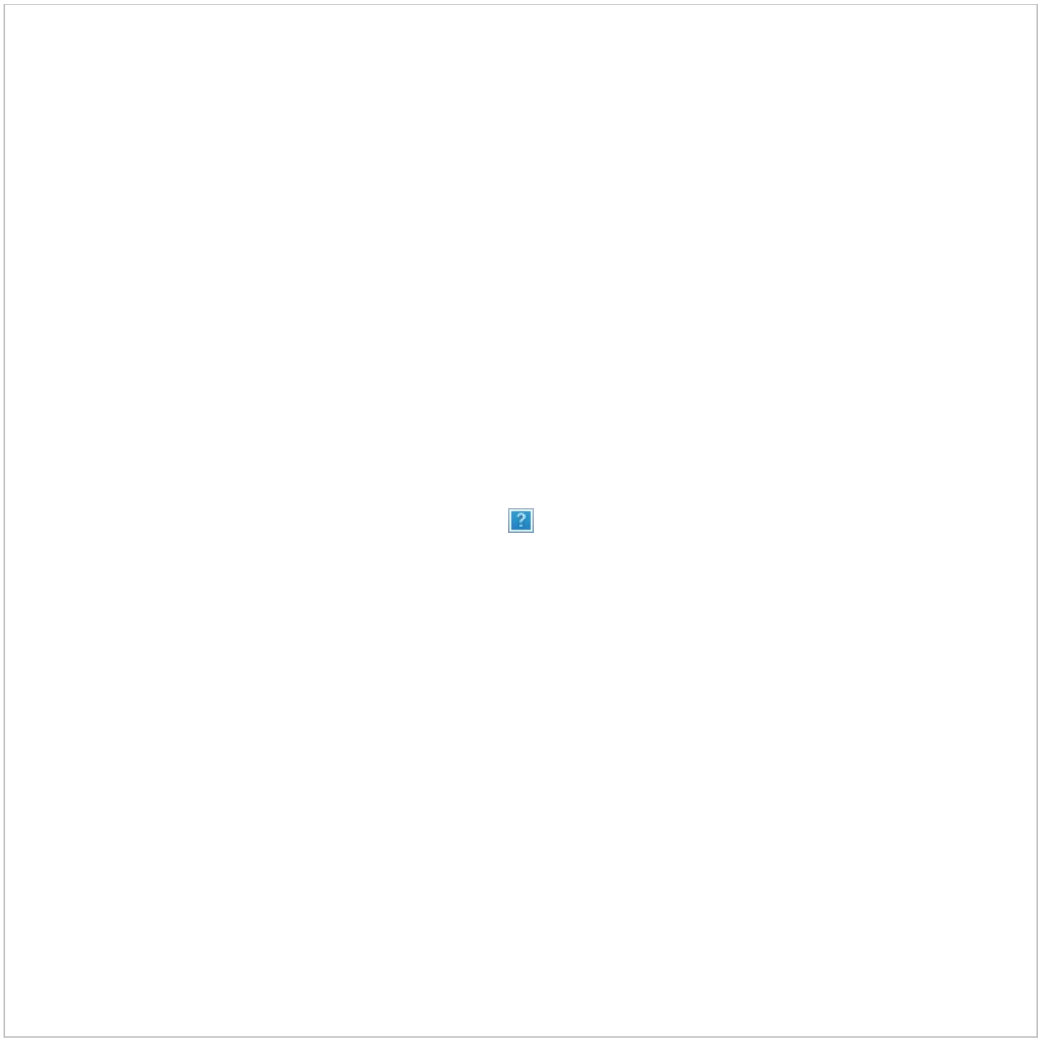
```
#Network Modularity = .23
House_13_14_opt <- cluster_optimal(House_13_14_network)
#House_13_14_opt_plot <- plot(House_13_14_opt, House_13_14_network, vertex.color = V(House_13_14_network)$color, vertex.size
        = 10) + title(main = "Network Modularity: ", sub = modularity(House_13_14_opt)) \
House_13_14_opt_plot <- plot(House_13_14_opt, House_13_14_network, vertex.label = NA, vertex.color = "black", edge.color =
        E(House_13_14_network)$color, vertex.size = 15) + title(main = "CA House 13-14 Network Modularity: ", sub =
        modularity(House_13_14_opt))
```



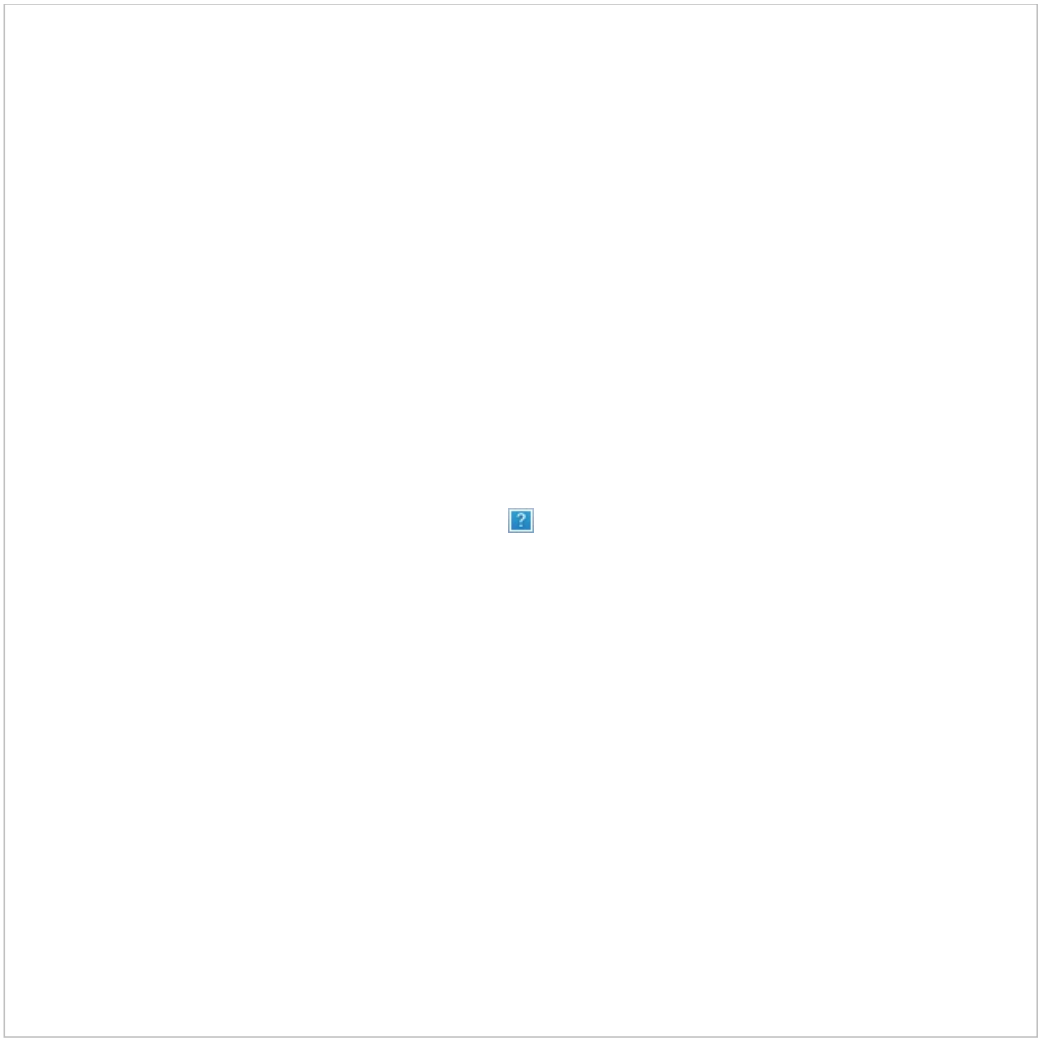
```
#Network Modularity = .33
House_15_16_opt <- cluster_optimal(House_15_16_network)
#House_15_16_opt_plot <- plot(House_15_16_opt, House_15_16_network, vertex.color = V(House_15_16_network)$color, vertex.size
        = 10) + title(main = "Network Modularity: ", sub = modularity(House_15_16_opt))
House_15_16_opt_plot <- plot(House_15_16_opt, House_15_16_network, vertex.label = NA, vertex.color = "black", edge.color =
        E(House_15_16_network)$color, vertex.size = 15) + title(main = "CA House 15-16 Network Modularity: ", sub =
        modularity(House_15_16_opt))
```



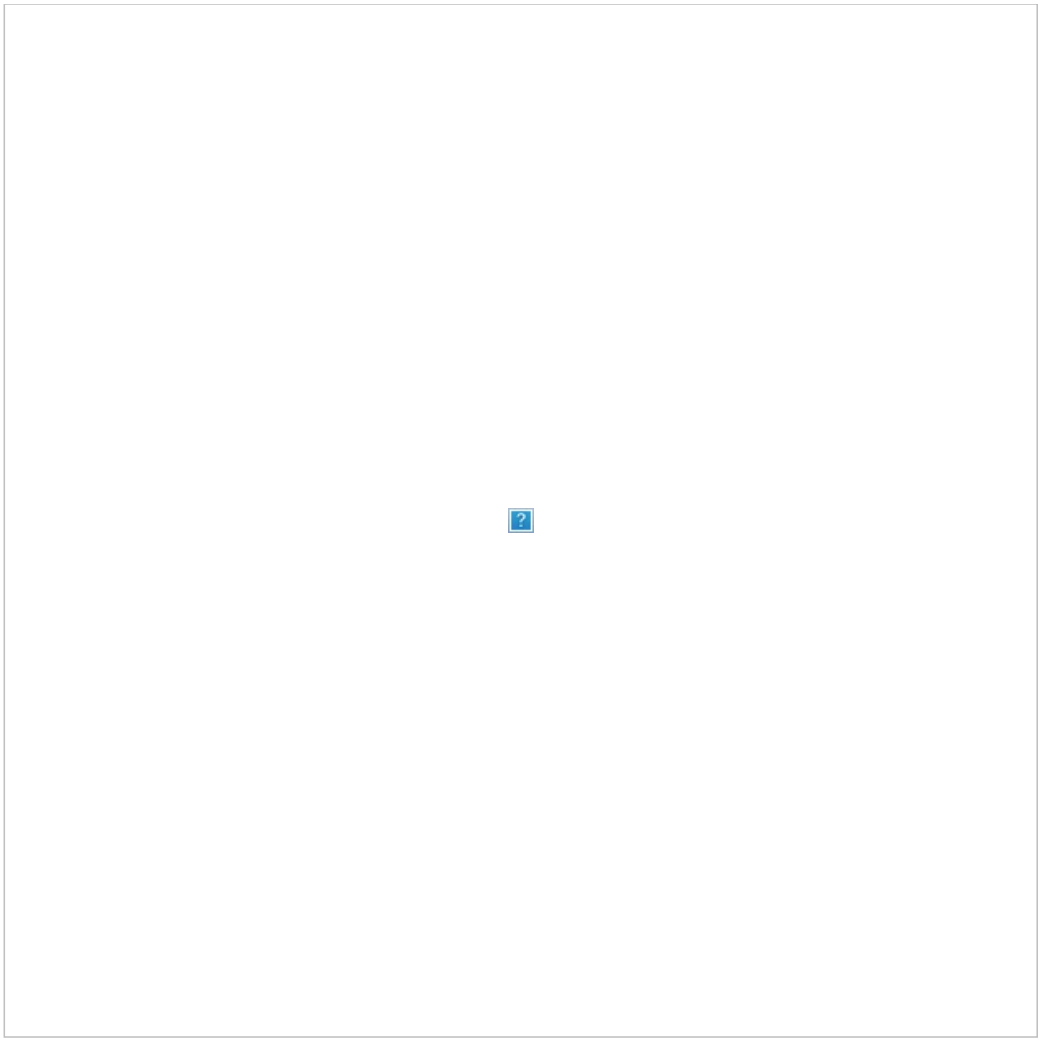
```
#Network Modularity increased every yearly cycle implying a higher level of polarization as time passed


#Community Detection Algorithm for Senate 2011-2016

#We want to see how the top echelon of political donors are grouped to detect any patterns

#Optimal Community Detection Algorithm to maximize modularity score increasing cohesiveness in the groups

#Network modularity to determine level of polarization per year especially among the more avid donors to see if they exhibit
        a tendency to only interact within their own party group

#Network Modularity = 0.2
Senate_11_12_opt <- cluster_optimal(Senate_11_12_network)
#Senate_11_12_opt_plot <- plot(Senate_11_12_opt, Senate_11_12_network, vertex.color = V(Senate_11_12_network)$color,
        vertex.size = 10) + title(main = "Network Modularity: ", sub = modularity(Senate_11_12_opt))
Senate_11_12_opt_plot <- plot(Senate_11_12_opt, Senate_11_12_network, vertex.label = NA, vertex.color = "black", edge.color =
        E(Senate_11_12_network)$color, vertex.size = 15) + title(main = "CA Senate 11-12 Network Modularity: ", sub =
        modularity(Senate_11_12_opt))
```



```
#Network Modularity = 0.15
Senate_13_14_opt <- cluster_optimal(Senate_13_14_network)
#Senate_13_14_opt_plot <- plot(Senate_13_14_opt, Senate_13_14_network, vertex.color = V(Senate_13_14_network)$color,
        vertex.size = 10) + title(main = "Network Modularity: ", sub = modularity(Senate_13_14_opt))
Senate_13_14_opt_plot <- plot(Senate_13_14_opt, Senate_13_14_network, vertex.label = NA, vertex.color = "black", edge.color =
        E(Senate_13_14_network)$color, vertex.size = 15) + title(main = "CA Senate 13-14 Network Modularity: ", sub =
        modularity(Senate_13_14_opt))
```



```
#Network Modularity = 0.22
Senate_15_16_opt <- cluster_optimal(Senate_15_16_network)
#Senate_15_16_opt_plot <- plot(Senate_15_16_opt, Senate_15_16_network, vertex.color = V(Senate_15_16_network)$color,
        vertex.size = 10) + title(main = "Network Modularity: ", sub = modularity(Senate_15_16_opt))
Senate_15_16_opt_plot <- plot(Senate_15_16_opt, Senate_15_16_network, vertex.label = NA, vertex.color = "black", edge.color =
        E(Senate_15_16_network)$color, vertex.size = 15) + title(main = "CA Senate 15-16 Network Modularity: ", sub =
        modularity(Senate_15_16_opt))
```



```
#The Network Modularity remained relatively stable with little change meaning there wasn't a significant increase in
        polarization
```

```
#function to isolate the industries or groups within the top 25% prevalence separated by party

isolate_both_groups <- function(df) {
  dem_group <- df %>%
    filter(PerDem >= 50) %>%
    select(EID, PerDem, CatCodeGroup) %>%
    rename(dem_donor = EID, dem_group = CatCodeGroup) %>%
    group_by(dem_group) %>%
    summarize(freq = n()) %>%
    mutate(freq_cat = ntile(freq, 4)) %>%
    filter(freq_cat == 4)

  rep_group <- df %>%
    filter(PerRep >= 50) %>%
    select(EID, PerRep, CatCodeGroup) %>%
    rename(rep_donor = EID, rep_group = CatCodeGroup) %>%
    group_by(rep_group) %>%
    summarize(freq = n()) %>%
    mutate(freq_cat = ntile(freq, 4)) %>%
    filter(freq_cat == 4 )

  both_groups <- dem_group %>%
   bind_rows(rep_group) %>%
   select(-freq_cat)

  return(both_groups)

}
```

```
#CA House (2011-2016) data and network

House_11_12_groups <- isolate_both_groups(CA_2011_2012_House)
House_13_14_groups <- isolate_both_groups(CA_2013_2014_House)
House_15_16_groups <- isolate_both_groups(CA_2015_2016_House)

House_11_12_groups
```

```
# A tibble: 8 × 3
  dem_group                          freq rep_group
  <chr>                             <int> <chr>
1 Finance, Insurance & Real Estate   120 <NA>
2 Health                             107 <NA>
3 Labor                              217 <NA>
4 Uncoded                            353 <NA>
5 <NA>                                32 Agriculture
6 <NA>                                46 Finance, Insurance & Real Estate
7 <NA>                                34 General Business
8 <NA>                               153 Uncoded
```

```
House_13_14_groups
```

```
# A tibble: 7 × 3
  dem_group                          freq rep_group
  <chr>                             <int> <chr>
1 Finance, Insurance & Real Estate   118 <NA>
2 Health                             119 <NA>
3 Labor                              201 <NA>
4 Uncoded                            415 <NA>
5 <NA>                                18 Energy & Natural Resources
6 <NA>                                19 Finance, Insurance & Real Estate
7 <NA>                               126 Uncoded
```

```
House_15_16_groups
```

```
# A tibble: 8 × 3
  dem_group                          freq rep_group
  <chr>                             <int> <chr>
1 Finance, Insurance & Real Estate    91 <NA>
2 Health                             102 <NA>
3 Labor                              154 <NA>
4 Uncoded                            352 <NA>
5 <NA>                                15 Agriculture
6 <NA>                                28 Finance, Insurance & Real Estate
7 <NA>                                22 General Business
8 <NA>                               118 Uncoded
```

```
#CA Senate (2011-2016) data and network

Senate_11_12_groups <- isolate_both_groups(CA_2011_2012_Senate)
Senate_13_14_groups <- isolate_both_groups(CA_2013_2014_Senate)
Senate_15_16_groups <- isolate_both_groups(CA_2015_2016_Senate)
```

```
Senate_11_12_groups
```

```
# A tibble: 7 × 3
  dem_group                          freq rep_group
  <chr>                             <int> <chr>
1 Finance, Insurance & Real Estate    58 <NA>
2 Health                              56 <NA>
3 Labor                              115 <NA>
4 Uncoded                             92 <NA>
5 <NA>                                65 Finance, Insurance & Real Estate
6 <NA>                                40 General Business
7 <NA>                                74 Uncoded
```

```
Senate_13_14_groups
```

```
# A tibble: 7 × 3
  dem_group                          freq rep_group
  <chr>                             <int> <chr>
1 Finance, Insurance & Real Estate    73 <NA>
2 Health                              92 <NA>
3 Labor                              114 <NA>
4 Uncoded                            128 <NA>
5 <NA>                                40 Agriculture
6 <NA>                                39 Finance, Insurance & Real Estate
7 <NA>                                85 Uncoded
```

```
Senate_15_16_groups
```

```
# A tibble: 6 × 3
  dem_group  freq rep_group
  <chr>     <int> <chr>
1 Health       89 <NA>
2 Labor       125 <NA>
3 Uncoded     158 <NA>
```

```
4 <NA>         12 Finance, Insurance & Real Estate
5 <NA>          9 General Business
6 <NA>         28 Uncoded
```

```
#functions to find the probabilities of the five groups we wanted to focus on in regards to how often either group backed the
        same candidate versus mainly having ties within the same group separated by both parties

dem_probability_matrix <- function(nodes, edges) {
  dem_nodes_2 <- nodes %>%
    filter(PerDem >= 50, (CatCodeGroup == "Health" | CatCodeGroup == "Labor" | CatCodeGroup == "Agriculture" | CatCodeGroup
        == "General Business" | CatCodeGroup == "Finance, Insurance & Real Estate")) %>%
    mutate(dem_node = EID) %>%
    mutate(dem_group = CatCodeGroup) %>%
    select(dem_node, dem_group)

  dem_nodes_1 <- nodes %>%
    filter(PerDem >= 50, (CatCodeGroup == "Health" | CatCodeGroup == "Labor" | CatCodeGroup == "Agriculture" | CatCodeGroup
        == "General Business" | CatCodeGroup == "Finance, Insurance & Real Estate")) %>%
    mutate(dem_node = EID) %>%
    mutate(dem_group = CatCodeGroup) %>%
    select(EID, dem_node, dem_group) %>%
    inner_join(edges, by = c("EID" = "node_1")) %>%
    inner_join(dem_nodes_2, by = c("node_2" = "dem_node")) %>%
    select(-EID, -edge) %>%
    mutate(dem_node_2 = node_2) %>%
    mutate(dem_group_1 = dem_group.x) %>%
    mutate(dem_group_2 = dem_group.y) %>%
    select(-dem_group.x, -dem_group.y, -node_2)

  d_p_matrix <- dcast(as.data.table(dem_nodes_1), dem_group_1 ~ dem_group_2)
  total <- sum(as.data.frame(d_p_matrix[1:5, 2:6]))

  d_proportion_matrix <- as.data.frame(d_p_matrix) %>%
    mutate(Agriculture = Agriculture/total, `General Business` = d_p_matrix$`General Business`/total, Health = Health/total,
        Labor = Labor/total, `Finance, Insurance & Real Estate` = d_p_matrix$`Finance, Insurance & Real Estate`/ total)

  return(d_proportion_matrix)
}

rep_probability_matrix <- function(nodes, edges) {
  rep_nodes_2 <- nodes %>%
    filter(PerRep >= 50, (CatCodeGroup == "Health" | CatCodeGroup == "Labor" | CatCodeGroup == "Agriculture" | CatCodeGroup
        == "General Business" | CatCodeGroup == "Finance, Insurance & Real Estate")) %>%
    mutate(rep_node = EID) %>%
    mutate(rep_group = CatCodeGroup) %>%
    select(rep_node, rep_group)

rep_nodes_1 <- nodes %>%
    filter(PerRep >= 50, (CatCodeGroup == "Health" | CatCodeGroup == "Labor" | CatCodeGroup == "Agriculture" | CatCodeGroup
        == "General Business" | CatCodeGroup == "Finance, Insurance & Real Estate")) %>%
    mutate(rep_node = EID) %>%
    mutate(rep_group = CatCodeGroup) %>%
    select(EID, rep_node, rep_group) %>%
    inner_join(edges, by = c("EID" = "node_1")) %>%
    inner_join(rep_nodes_2, by = c("node_2" = "rep_node")) %>%
    select(-EID, -edge) %>%
    mutate(rep_node_2 = node_2) %>%
    mutate(rep_group_1 = rep_group.x) %>%
    mutate(rep_group_2 = rep_group.y) %>%
    select(-rep_group.x, -rep_group.y, -node_2)

  r_p_matrix <- dcast(as.data.table(rep_nodes_1), rep_group_1 ~ rep_group_2)
  total <- sum(as.data.frame(r_p_matrix[1:5, 2:6]))

  r_proportion_matrix <- as.data.frame(r_p_matrix) %>%
    mutate(Agriculture = Agriculture/total, `General Business` = r_p_matrix$`General Business`/total, Health = Health/total,
        Labor = Labor/total, `Finance, Insurance & Real Estate` = r_p_matrix$`Finance, Insurance & Real Estate`/ total)

  return(r_proportion_matrix)

}

#CA Senate 2015-16 has a different shape due to NO Agriculture donors prevalent who donated over 50% of their funds for
        Republican candidates

rep_nodes_2 <- CA_2015_2016_Senate %>%
    filter(PerRep >= 50, (CatCodeGroup == "Health" | CatCodeGroup == "Labor" | CatCodeGroup == "Agriculture" | CatCodeGroup
        == "General Business" | CatCodeGroup == "Finance, Insurance & Real Estate")) %>%
    mutate(rep_node = EID) %>%
    mutate(rep_group = CatCodeGroup) %>%
    select(rep_node, rep_group)

rep_nodes_1 <- CA_2015_2016_Senate %>%
    filter(PerRep >= 50, (CatCodeGroup == "Health" | CatCodeGroup == "Labor" | CatCodeGroup == "Agriculture" | CatCodeGroup
        == "General Business" | CatCodeGroup == "Finance, Insurance & Real Estate")) %>%
```

```
    mutate(rep_node = EID) %>%
    mutate(rep_group = CatCodeGroup) %>%
    select(EID, rep_node, rep_group) %>%
    inner_join(CA_2015_2016_Senate_edges, by = c("EID" = "node_1")) %>%
    inner_join(rep_nodes_2, by = c("node_2" = "rep_node")) %>%
    select(-EID, -edge) %>%
    mutate(rep_node_2 = node_2) %>%
    mutate(rep_group_1 = rep_group.x) %>%
    mutate(rep_group_2 = rep_group.y) %>%
    select(-rep_group.x, -rep_group.y, -node_2)

  r_p_matrix <- dcast(as.data.table(rep_nodes_1), rep_group_1 ~ rep_group_2)
  total <- sum(as.data.frame(r_p_matrix[1:4, 2:5]))

  r_proportion_matrix <- as.data.frame(r_p_matrix) %>%
    mutate(`General Business` = r_p_matrix$`General Business`/total, Health = Health/total, Labor = Labor/total, `Finance,
        Insurance & Real Estate` = r_p_matrix$`Finance, Insurance & Real Estate`/ total)


#Since the order of the tables made it so that Health-Labor and Labor-Health had different values this function sums up all
        the repeat pairing to give the true ranking of which industry pairings were more frequent for either party

all_ties_p <- function(p) {
  titles <- c("Finance-Agriculture", "Business-Agriculture", "Health-Agriculture", "Labor-Agriculture", "Finance-Business",
        "Finance-Health", "Finance-Labor", "Business-Health", "Business-Labor", "Health-Labor", "Agriculture-Agriculture",
        "Finance-Finance", "Business-Business", "Health-Health", "Labor-Labor")

  results <- c(sum(p[2, 2], p[1, 3]), sum(p[3, 2], p[1, 4]), sum(p[4, 2], p[1, 5]), sum(p[5, 2], p[1, 6]), sum(p[2, 4], p[3,
        3]), sum(p[2, 5], p[4, 3]), sum(p[2, 6], p[5, 3]), sum(p[3, 5], p[4, 4]), sum(p[3, 6], p[5, 4]), sum(p[4, 6], p[5,
        5]), p[1, 2], p[2, 3], p[3, 4], p[4, 5], p[5, 6])

  total <- as.data.frame(results, row.names = titles)

  total <- total %>%
    arrange(desc(results))

  return(total)
}

#The remaining code chunks output the probability matrices for all House and Senate years (2011-16) separated by party

all_ties_p(dem_probability_matrix(CA_2011_2012_House, CA_2011_2012_House_edges))
```

```
                          results
Labor-Labor             0.128457843
Finance-Health          0.122701673
Health-Labor            0.105122577
Finance-Labor           0.101519795
Finance-Business        0.095494451
Business-Health         0.092015902
Business-Labor          0.077314892
Health-Health           0.067852410
Finance-Finance         0.065533378
Business-Business       0.037560046
Finance-Agriculture     0.028573795
Health-Agriculture      0.027931920
Business-Agriculture    0.022983270
Labor-Agriculture       0.022486334
Agriculture-Agriculture 0.004451714
```

```
all_ties_p(rep_probability_matrix(CA_2011_2012_House, CA_2011_2012_House_edges))
```

```
                          results
Finance-Business        0.157242213
Finance-Agriculture     0.135470215
Business-Agriculture    0.124281826
Agriculture-Agriculture 0.105836105
Finance-Health          0.098881161
Business-Health         0.091926217
Finance-Finance         0.085576051
Business-Business       0.065618385
Health-Agriculture      0.061687330
Health-Health           0.028424554
Business-Labor          0.013607499
Finance-Labor           0.011793166
Labor-Agriculture       0.009978833
Health-Labor            0.007559722
Labor-Labor             0.002116722
```

```
all_ties_p(dem_probability_matrix(CA_2013_2014_House, CA_2013_2014_House_edges))
```

```
                          results
Health-Labor            0.11672593
Finance-Health          0.10811654
Labor-Labor             0.10549629
Finance-Labor           0.10063011
Business-Health         0.09248861
Finance-Business        0.09178676
Business-Labor          0.09050783
Health-Health           0.06349429
Finance-Finance         0.05229584
Business-Business       0.03919458
Health-Agriculture      0.03554495
Finance-Agriculture     0.03552935
Business-Agriculture    0.03247239
Labor-Agriculture       0.02751263
Agriculture-Agriculture 0.00820388
```

```
all_ties_p(rep_probability_matrix(CA_2013_2014_House, CA_2013_2014_House_edges))
```

```
                          results
Business-Agriculture    0.164887307
Finance-Business        0.143534994
Finance-Agriculture     0.106761566
Agriculture-Agriculture 0.088967972
Finance-Health          0.087781732
Business-Health         0.084223013
Finance-Finance         0.080664294
Health-Agriculture      0.071174377
Business-Business       0.067615658
Health-Health           0.030842230
Business-Labor          0.022538553
Labor-Agriculture       0.018979834
Health-Labor            0.015421115
Finance-Labor           0.010676157
Labor-Labor             0.005931198
```

```
all_ties_p(dem_probability_matrix(CA_2015_2016_House, CA_2015_2016_House_edges))
```

```
                          results
Health-Labor            0.133880507
Finance-Health          0.109474955
Finance-Labor           0.104477972
Labor-Labor             0.101750151
Business-Health         0.091852746
Finance-Business        0.083114062
Business-Labor          0.082510561
Health-Health           0.075051298
Finance-Finance         0.051828606
Business-Business       0.036330718
Health-Agriculture      0.034351237
Finance-Agriculture     0.032130356
Business-Agriculture    0.029668075
Labor-Agriculture       0.027012674
Agriculture-Agriculture 0.006566083
```

```
all_ties_p(rep_probability_matrix(CA_2015_2016_House, CA_2015_2016_House_edges))
```

```
                          results
Finance-Business        0.194070081
Business-Agriculture    0.143755615
Finance-Agriculture     0.118598383
Finance-Finance         0.117699910
Finance-Health          0.093441150
Business-Health         0.088948787
Business-Business       0.081761006
Agriculture-Agriculture 0.061994609
Health-Agriculture      0.054806828
Health-Health           0.021563342
Labor-Agriculture       0.009883199
Business-Labor          0.008086253
Health-Labor            0.003593890
Finance-Labor           0.001796945
Labor-Labor             0.000000000
```

```
all_ties_p(dem_probability_matrix(CA_2011_2012_Senate, CA_2011_2012_Senate_edges))
```

```
                          results
Labor-Labor             0.159932156
Health-Labor            0.138281952
Finance-Labor           0.137783099
```

```
Business-Labor          0.111393794
Finance-Health          0.087299212
Finance-Business        0.079068143
Business-Health         0.066447172
Finance-Finance         0.049835379
Health-Health           0.041554425
Labor-Agriculture       0.031677143
Business-Business       0.029332535
Finance-Agriculture     0.024593435
Business-Agriculture    0.019854335
Health-Agriculture      0.019804450
Agriculture-Agriculture 0.003142772
```

```
all_ties_p(rep_probability_matrix(CA_2011_2012_Senate, CA_2011_2012_Senate_edges))
```

```
                           results
Finance-Business        0.1754471704
Finance-Health          0.1513048578
Finance-Finance         0.1334180432
Finance-Agriculture     0.1207115629
Business-Health         0.0969602189
Business-Agriculture    0.0778027563
Health-Agriculture      0.0667578927
Business-Business       0.0561039977
Health-Health           0.0473072036
Agriculture-Agriculture 0.0290294204
Finance-Labor           0.0168116509
Business-Labor          0.0110448636
Health-Labor            0.0101651842
Labor-Agriculture       0.0067442088
Labor-Labor             0.0003909686
```

```
all_ties_p(dem_probability_matrix(CA_2013_2014_Senate, CA_2013_2014_Senate_edges))
```

```
                          results
Health-Labor            0.145265516
Finance-Health          0.137827048
Finance-Labor           0.113357650
Labor-Labor             0.104569344
Health-Health           0.097475516
Business-Health         0.094632241
Business-Labor          0.092909044
Finance-Business        0.073264597
Finance-Finance         0.052815991
Business-Business       0.025532037
Health-Agriculture      0.019070048
Labor-Agriculture       0.017059651
Finance-Agriculture     0.014302536
Business-Agriculture    0.010884862
Agriculture-Agriculture 0.001033918
```

```
all_ties_p(rep_probability_matrix(CA_2013_2014_Senate, CA_2013_2014_Senate_edges))
```

```
                           results
Business-Agriculture    0.1846130079
Finance-Agriculture     0.1809052989
Finance-Business        0.1569596787
Agriculture-Agriculture 0.1154024409
Finance-Finance         0.0756990576
Business-Business       0.0750811061
Health-Agriculture      0.0665842731
Business-Health         0.0554611463
Finance-Health          0.0548431948
Health-Health           0.0095782481
Finance-Labor           0.0078788815
Business-Labor          0.0075699058
Labor-Agriculture       0.0067974664
Health-Labor            0.0024718060
Labor-Labor             0.0001544879
```

```
all_ties_p(dem_probability_matrix(CA_2015_2016_Senate, CA_2015_2016_Senate_edges))
```

```
                          results
Health-Labor            0.117718792
Finance-Health          0.109609365
Finance-Labor           0.101723507
Business-Health         0.095687167
Business-Labor          0.088146823
Finance-Business        0.087211902
Labor-Labor             0.077618796
```

```
Health-Health           0.062639730
Finance-Finance         0.050140238
Health-Agriculture      0.045242063
Finance-Agriculture     0.042112109
Labor-Agriculture       0.038514695
Business-Business       0.037356205
Business-Agriculture    0.037091988
Agriculture-Agriculture 0.009186618
```

```r
#2015-2016 Senate has no occurrence of any Agriculture group
r_titles <- c("Finance-Finance", "Business-Business", "Health-Health", "Labor-Labor", "Finance-Business", "Finance-Health",
        "Finance-Labor", "Business-Health", "Business-Labor", "Health-Labor")

r_results <- c(r_proportion_matrix[1, 2], r_proportion_matrix[2, 3], r_proportion_matrix[3, 4], r_proportion_matrix[4, 5],
        sum(r_proportion_matrix[1, 3], r_proportion_matrix[2, 3]), sum(r_proportion_matrix[1, 4], r_proportion_matrix[3, 2]),
        sum(r_proportion_matrix[1, 5], r_proportion_matrix[4, 2]), sum(r_proportion_matrix[2, 4], r_proportion_matrix[3, 3]),
        sum(r_proportion_matrix[2, 5], r_proportion_matrix[4, 3]), sum(r_proportion_matrix[3, 5], r_proportion_matrix[4, 4]))

r_total <- as.data.frame(r_results, row.names = r_titles)

r_total <- r_total %>%
   arrange(desc(r_results))

r_total
```

```
                   r_results
Finance-Business  0.333333333
Finance-Finance   0.212560386
Business-Business 0.140096618
Finance-Health    0.091787440
Business-Health   0.091787440
Business-Labor    0.028985507
Finance-Labor     0.019323671
Health-Health     0.009661836
Labor-Labor       0.000000000
Health-Labor      0.000000000
```

References Consulted

Reuning, Kevin. *Election Donation Networks*. Harvard Dataverse, 4 June 2020, https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi%3A10.7910%2FDVN%2FYMDFPW.

“Create Similarity Matrix.” Stack Overflow, 29 July 2018, https://stackoverflow.com/questions/52078017/create-similarity-matrix.