UN Votes Group Project

Group 10: Chew Shi Zhan A0216042W, Lily Rozana Joehann Aung A0239463X, Joey Yang Yixuan A02

2023-08-30

Contents

Introduction

The UN Votes TidyTuesday data contains three main datasets. The datasets are "unvotes", "roll_calls" and "issues". For "unvotes", it contains the voting result of a wide variety of countries (yes, no or abstain) classified according to their roll call id (rcid). For "roll_calls", it is a much wider dataset with more variables and information such as session number, "importantvote", indicating whether the vote was classified as important by the U.S. State Department which began only after the 39th session, as well as "short" and "long" that provide descriptions with the extent of their depth based on their title. They also have other essential details such as the date, resolution code, whether the vote was on an amendment and only on a paragraph and not a resolution. Lastly, the "issues" dataset succinctly describes the specific conflict happening, if there was, according to the specific roll call id which comes with an issue code. These three datasets can be easily joined together for a clearer overview with the column "rcid" that they have in common.

Descriptive Statistics

```r
# Tidying the data for Qn1
## Filtering out the data
unvotes_rdy <- unvotes %>%
  select(-c(country_code)) %>%
  left_join(roll_calls[, c("rcid", "date", "amend", "para")], by="rcid") %>%
  filter(format(date, format = "%Y") <= 1984,format(date, format = "%Y") >= 1956 ) %>%
  filter(para == 0) %>%
  mutate(vote = case_when(vote == "yes" ~ 1,
                          vote == "no" ~ -1,
                          vote == "abstain" ~ 0.5)) %>%
  select(rcid, country, vote) %>%
  pivot_wider(names_from=rcid, values_from=vote) %>%
  replace(is.na(.), 0)

## Plotting the cosine similarity matrix
voting_matrix <- unvotes_rdy[-1]
similarity_matrix <- cosine(t(as.matrix(voting_matrix)))
rownames(similarity_matrix) <- unvotes_rdy$country
colnames(similarity_matrix) <- unvotes_rdy$country
similarity_df <- as.data.frame(similarity_matrix)

# Descriptive Statistics for Qn1
a <- similarity_df %>%
  mutate_all(median)
b <- similarity_df %>%
  mutate_all(min)
c <- rbind(a[1,], b[1,])
rownames(c) <- c("Median of Similarity", "Min of Similarity")
c
```

```
##                      United States     Canada       Cuba      Haiti
## Median of Similarity    0.06312303 0.50510518  0.6809475 0.67727824
## Min of Similarity      -0.24520754 0.02600255 -0.1182627 0.01327681
##                      Dominican Republic     Mexico  Guatemala  Honduras
## Median of Similarity         0.73396876 0.81601039 0.66463121 0.7050628
## Min of Similarity            0.02534897 0.02360632 0.02631807 0.0255655
##                      El Salvador  Nicaragua Costa Rica    Panama  Colombia
## Median of Similarity  0.72986269 0.67689055 0.72023616 0.7853080 0.7799732
## Min of Similarity     0.02541438 0.02517566 0.02463884 0.0120666 0.0121366
##                      Venezuela    Ecuador       Peru    Brazil    Bolivia
## Median of Similarity 0.8074592 0.80547196 0.80354598 0.7629783 0.74059880
## Min of Similarity    0.0237641 0.02422508 0.01191574 0.0240859 0.02504897
##                        Paraguay      Chile  Argentina    Uruguay United Kingdom
## Median of Similarity 0.61409023 0.74062171 0.78160091 0.73507055     0.32443655
## Min of Similarity    0.02769244 0.02437394 0.02411214 0.02449429    -0.04968581
##                         Ireland Netherlands    Belgium Luxembourg     France
## Median of Similarity 0.67889788  0.53149525 0.47226036 0.47216438 0.41856882
## Min of Similarity    0.02550514  0.02496676 0.02554039 0.02554039 0.01387818
##                           Spain   Portugal     Poland    Austria    Hungary
## Median of Similarity 0.71671793 0.51321446  0.5687544 0.74551908  0.5589421
## Min of Similarity    0.01267449 0.02628169 -0.2085280 0.02591944 -0.2147385
```

```
##                       Czechoslovakia      Italy    Albania  Yugoslavia
## Median of Similarity      0.5566448 0.53487795  0.4131181  0.77148874
## Min of Similarity        -0.2128542 0.02652091 -0.2802788 -0.09838992
##                          Greece   Bulgaria    Romania     Russia    Ukraine
## Median of Similarity 0.76442607  0.5564420  0.7241712  0.5544746  0.5532166
## Min of Similarity    0.02461271 -0.2136283 -0.1863507 -0.2162324 -0.2167445
##                         Belarus    Finland     Sweden     Norway    Denmark
## Median of Similarity  0.5561418 0.73955891 0.71897544 0.62751450 0.61649302
## Min of Similarity    -0.2161595 0.02560738 0.02554672 0.02543492 0.02568327
##                         Iceland    Liberia    Ethiopia South Africa       Libya
## Median of Similarity 0.62116502 0.77512206  0.77397214  0.006332489  0.74758615
## Min of Similarity    0.02577054 0.02445214 -0.02990646 -0.280278797 -0.04345204
##                            Iran     Turkey       Iraq       Egypt      Syria
## Median of Similarity 0.79100154 0.72928532  0.7338758  0.77274115  0.7318325
## Min of Similarity    0.02383995 0.02402847 -0.1223717 -0.09049416 -0.1000228
##                         Lebanon    Jordan     Israel Saudi Arabia
## Median of Similarity 0.75183297 0.8004487  0.2716102    0.7668975
## Min of Similarity    0.01378664 0.0120073 -0.1712424   -0.0436331
##                      Yemen Arab Republic Afghanistan      Taiwan       India
## Median of Similarity           0.7459354  0.73969624  0.20450215  0.76574813
## Min of Similarity            -0.0731956 -0.05306043 -0.05317888 -0.05895266
##                        Pakistan Myanmar (Burma)   Sri Lanka       Nepal
## Median of Similarity 0.78636038      0.79162102  0.78366485  0.80428250
## Min of Similarity    0.02353593     -0.02923487 -0.02939173 -0.03226394
##                        Thailand     Cambodia      Laos Philippines   Indonesia
## Median of Similarity 0.81099901  0.663825397 0.7389462  0.80132670  0.78962640
## Min of Similarity    0.02363102 -0.009391182 0.0000000  0.02351152 -0.03887367
##                       Australia New Zealand       Sudan     Morocco     Tunisia
## Median of Similarity 0.59062101   0.6083745  0.76958803  0.76766198  0.80799287
## Min of Similarity    0.02573638   0.0255927 -0.09420799 -0.07657989 -0.02131157
##                           Japan       Ghana   Malaysia     Guinea     Cyprus
## Median of Similarity 0.64674937  0.79808416 0.81611694  0.7414624 0.80414627
## Min of Similarity    0.02607316 -0.04172335 0.02367735 -0.1313835 0.02445214
##                            Mali      Senegal      Benin      Niger
## Median of Similarity  0.7693538  0.799869850 0.78344286 0.79184090
## Min of Similarity    -0.1132094 -0.004243847 0.02508638 0.02480695
##                      Côte d'Ivoire Burkina Faso       Togo   Cameroon     Gabon
## Median of Similarity    0.79217507    0.7844262 0.80603451 0.80383109 0.7753217
## Min of Similarity       0.02569175    0.0000000 0.00761796 0.01477154 0.0000000
##                      Central African Republic       Chad Congo - Brazzaville
## Median of Similarity               0.73786485 0.77704117          0.73407064
## Min of Similarity                  0.01460862 0.02502159         -0.03868218
##                      Madagascar     Nigeria     Somalia Congo - Kinshasa
## Median of Similarity 0.79585773  0.80310941  0.75802219        0.7847008
## Min of Similarity    0.02420201 -0.04437599 -0.07523608        0.0239642
##                      Sierra Leone   Mongolia  Mauritania    Tanzania    Jamaica
## Median of Similarity   0.78990875  0.5933457  0.77522529  0.78314849 0.80109755
## Min of Similarity     -0.04749864 -0.1526942 -0.06302972 -0.07473257 0.01292425
##                          Burundi     Rwanda Trinidad & Tobago     Algeria
## Median of Similarity  0.79255249 0.79279475       0.803298385  0.77769783
## Min of Similarity    -0.04328163 0.02520163      -0.002046459 -0.07709395
##                          Uganda      Kuwait       Kenya    Zanzibar     Malta
## Median of Similarity  0.79103676  0.79267650  0.80017211  0.01327681 0.7650394
## Min of Similarity    -0.03953036 -0.02342166 -0.02574535 -0.03034136 0.0000000
```

```
##                          Zambia    Malawi    Maldives    Singapore      Gambia
## Median of Similarity  0.79157408 0.7018323  0.74622259  0.795469214  0.68321817
## Min of Similarity    -0.05413419 0.0000000 -0.01194102 -0.003941311 -0.01976655
##                          Guyana   Lesotho     Botswana Barbados
## Median of Similarity  0.79507741 0.7582174  0.735097591 0.763115
## Min of Similarity    -0.01405827 0.0000000 -0.001311824 0.000000
##                      Yemen People's Republic    Mauritius Equatorial Guinea
## Median of Similarity              0.75292803  0.751332881         0.6143325
## Min of Similarity                -0.08344211 -0.001933966        -0.0334193
##                       Eswatini      Fiji      Bhutan      Bahrain     Qatar
## Median of Similarity 0.7142995 0.7693564  0.77440350  0.799722434 0.7947753
## Min of Similarity    0.0000000 0.0000000 -0.01219037 -0.002253808 0.0000000
##                             Oman       China United Arab Emirates
## Median of Similarity  0.784273695  0.69280333          0.792581767
## Min of Similarity    -0.001234355 -0.08892314         -0.007552931
##                      Federal Republic of Germany German Democratic Republic
## Median of Similarity                   0.4098975                  0.5671886
## Min of Similarity                      0.0000000                 -0.1536474
##                          Bahamas  Bangladesh     Grenada Guinea-Bissau
## Median of Similarity  0.727550628  0.73492214  0.53653227    0.69417803
## Min of Similarity    -0.008096436 -0.04183538 -0.08068715   -0.07871325
##                      Cape Verde São Tomé & Príncipe Mozambique     Comoros
## Median of Similarity  0.6561674           0.6886210  0.6854300  0.53964420
## Min of Similarity    -0.1360805          -0.1384366 -0.1601212 -0.01702585
##                      Papua New Guinea    Suriname     Angola   Djibouti
## Median of Similarity       0.69940239  0.70673694  0.6312594  0.6596701
## Min of Similarity         -0.03379403 -0.05211031 -0.1883171 -0.1076849
##                        Vietnam       Samoa Seychelles Solomon Islands
## Median of Similarity  0.6049107  0.61835736  0.4516516      0.34403179
## Min of Similarity    -0.2144149 -0.02431053 -0.1997435     -0.08054261
##                        St. Lucia   Zimbabwe St. Vincent & Grenadines    Vanuatu
## Median of Similarity  0.48515116  0.4906998                0.3084751  0.4467491
## Min of Similarity    -0.08049931 -0.1974003               -0.1134478 -0.1629910
##                          Belize Antigua & Barbuda   Dominica St. Kitts & Nevis
## Median of Similarity  0.3788094         0.3588843  0.2209078        0.11232396
## Min of Similarity    -0.1186796        -0.1076488 -0.0623587       -0.07454993
##                         Brunei
## Median of Similarity  0.2763771
## Min of Similarity    -0.1308215
```

We first tidied the data by filtering out the years after and inclusive of 1985 as these years do not have values under the `para` column, and before and inclusive of 1955 to ensure both voting blocs have been established by then. The `para` column is essential as it allows us to only count votes that are on full resolutions. As highlighted in a report to the US Congress, including preliminary votes in calculating voting coincidence tends to distort the metric as these votes may not correspond to the country's final stance on the resolution. We thus keep only the rows where `para` equals 0. We then assign "yes", "no", "abstain" votes with values 1, -1, and 0.5 respectively. NA values were created when pivoting the data wider, as some countries were not UN member states when some resolutions were voted on. These were replaced with the value 0. After cleaning the data, we calculated the cosine similarity between the voting patterns of countries and calculated the median and minimum value of pairwise similarity of each country against the other countries. Maximum value is not informative as it is always 1 since the country would be compared against itself. For instance then, the United States' median similarity to other countries is about 0, suggesting it is generally not similar to other countries.

```r
# Tidying the data for Qn2
## Filtering the important votes
combined <- unvotes %>%
  left_join(issues,by='rcid') %>%
  left_join(roll_calls,by='rcid') %>%
  select(rcid,country,vote,issue,short,session,importantvote,date) %>%
  filter(importantvote==1& !is.na(issue))

## Calculate number of votes and the similarity index of each issue
important <- combined %>%
  group_by(rcid,issue,date,vote) %>%
  summarise(pro = n()) %>%
  spread(key=vote,value=pro) %>%
  mutate(similarity_index = (abs(yes-no)/(yes+no))*100) %>%
  drop_na(issue) %>%
  arrange(similarity_index)

# Descriptive Statistics for Qn2
## Distribution of Similarity Index
lapply(split(important$similarity_index,important$issue), summary)
```

```
## $`Arms control and disarmament`
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   10.00   71.44   95.00   80.74   98.70   98.87      13
##
## $Colonialism
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   60.00   80.85   84.47   85.22   94.59   97.69       2
##
## $`Economic development`
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   44.83   91.30   96.52   90.24   97.89   98.87       5
##
## $`Human rights`
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   8.219  49.541  81.241  69.659  87.096  98.889       3
##
## $`Nuclear weapons and nuclear material`
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   10.00   79.03   96.18   82.53   98.75   98.91       2
##
## $`Palestinian conflict`
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   10.00   82.24   86.09   84.04   90.33   98.63       4
```

We also joined the three data sets and filter out the rows with the important vote entries as 1 in order to filter out the issues that are deemed as important by the US. We then counted the number of votes for each type of vote then spread the votes into three separate columns to show the count for each type of vote for each date and issue. The yes and no vote counts are used in the calculation of the similarity indexes. In order to obtain the more polarised issues, the rows with similarity index less than 50 are filtered out. The occurrence of each issue is tallied and the highest count obtained was for human rights issues. The top 10 human rights issues with the lowest similarity index are filtered out and the short descriptions are converted to lowercase. The summary statistics of the similarity index shows that human rights is one issue that seems to be extremely polarising as compared to the other issues, with its 25th percentile being at 49.79 as compared to those in the other issues, which is all above 80.

Question 1: Voting Patterns Amongst Countries

Introduction

Our first research question would be "What countries seem to be voting in a synchronised/similar or opposing pattern and are there indicators of alliance via voting blocs or hostility?". We were keen to investigate if there was any correlation between the voting pattern of countries within certain blocs associated with the UN. Analysing the data would potentially help us to understand geopolitical dynamics between countries. Interpreting the pattern of votes between individual countries as well as between blocs of countries reveals similarity in the foreign policy beliefs that underlie voting decisions.

Methodology

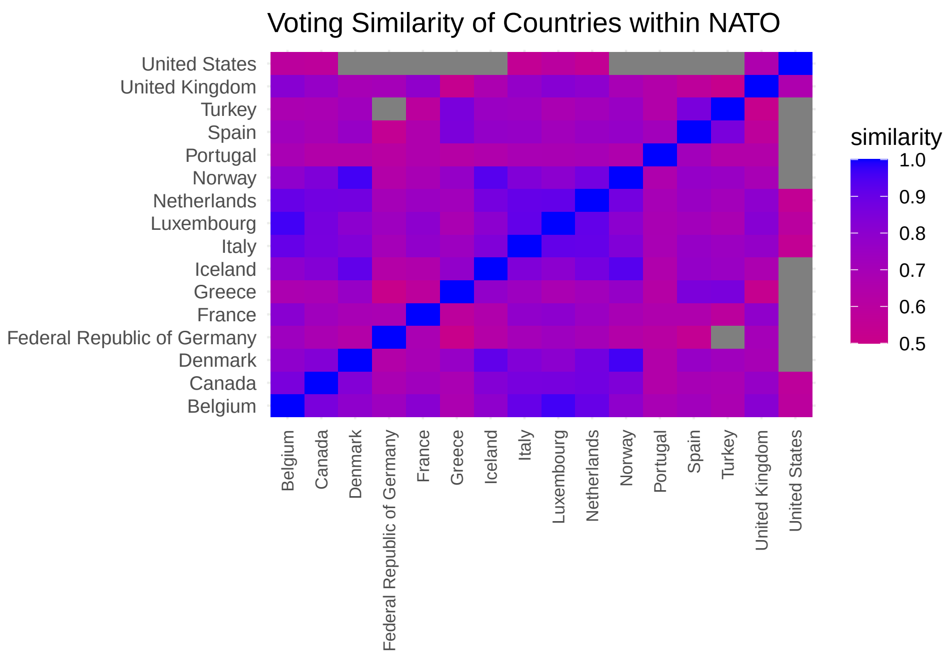
Using the similarity matrix, we created two plots. First, a heatmap to answer question 1 on a more granular level. The heatmap plots the historical voting similarity between members of within NATO, within the Warsaw Pact and between select members across the two alliances. The heatmap is highly appropriate for visualising this information as the information is pairwise and there are a limited number of countries to plot for each group. The correspondence of colour intensity to the similarity scale also allows the viewer to quickly glean which countries are voting similarly. We focus on NATO and the Warsaw Pact as the major voting blocs as the period our dataset covers overlaps with the Cold War.

Next, we created a boxplot that represents the voting pattern within the groups of the countries. We plotted their similarities within their individual groups. This method is efficient in analysing the bigger picture of the voting patterns. For instance, we see the specific range as well as inter-quartile range of the similarities within each bloc, reflecting the spread and polarity of the votes. This representation also clearly allows us to compare the median of similarities within groups in a side-by-side manner. This can directly help us to answer the question about how similar the voting patterns are within blocs and help us conclude if it shows signs of alliances compared to other blocs. As it is better suited for presenting summarised data, we are able to explore the voting similarity of the larger 19 member Non-Aligned Movement (NAM) group in this plot.

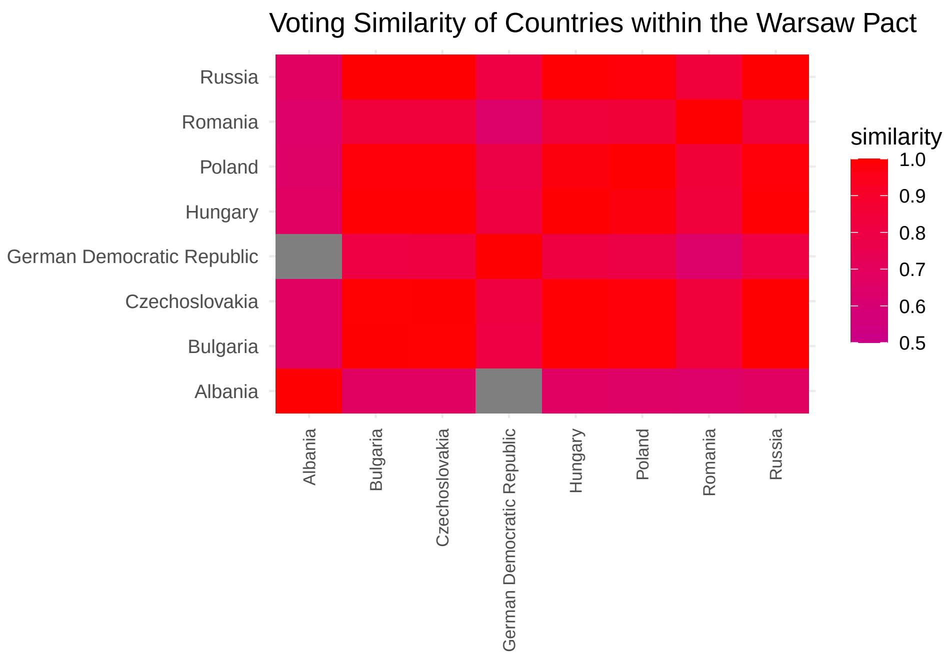
Visualisation

Plot 1

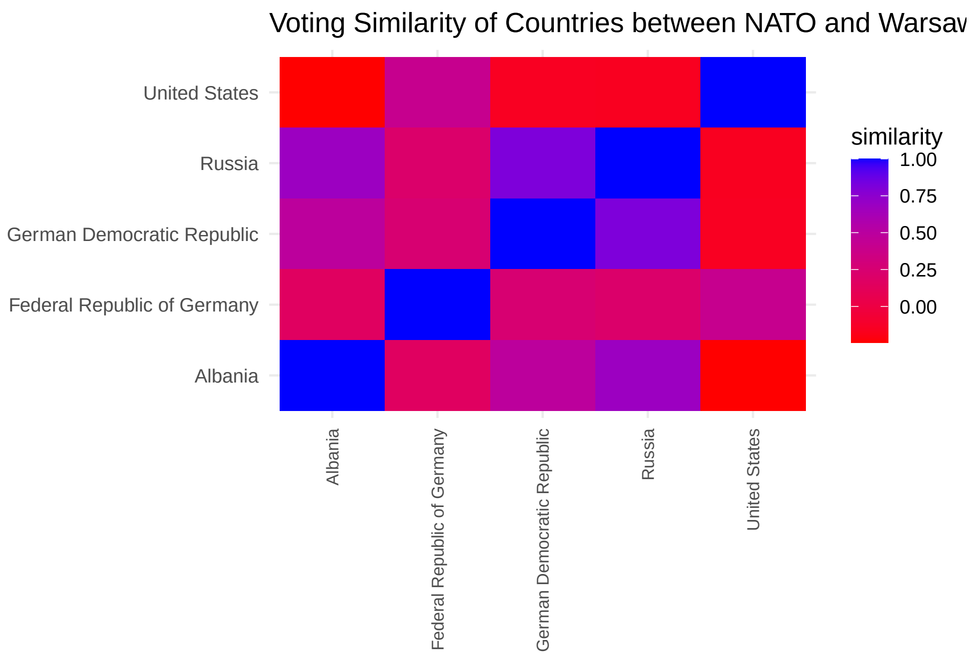
```r
NATO <- c("Belgium", "Canada", "Denmark", "France", "Iceland", "Italy", "Luxembourg", "Netherlands", "No

NATO_tidy_matrix <- similarity_matrix[NATO, NATO] %>%
  as_tibble(rownames = "country1") %>%
  pivot_longer(-country1, names_to = "country2", values_to = "similarity")

ggplot(NATO_tidy_matrix, aes(x = country1, y = country2, fill = similarity)) +
  geom_tile() +
  scale_fill_gradient2(low = "white", mid = "red", high = "blue", limits = c(0.5, 1.0)) +
  labs(title = "Voting Similarity of Countries within NATO", x = "", y = "") +
  theme_minimal() +
  theme(axis.text.x  = element_text(angle=90, vjust=0.5, size=8, hjust=1))
```




```
Warsaw_Pact <- c("Albania", "Bulgaria", "Czechoslovakia", "German Democratic Republic", "Hungary", "Pola

WP_tidy_matrix <- similarity_matrix[Warsaw_Pact, Warsaw_Pact] %>%
  as_tibble(rownames = "country1") %>%
  pivot_longer(-country1, names_to = "country2", values_to = "similarity")

ggplot(WP_tidy_matrix, aes(x = country1, y = country2, fill = similarity)) +
  geom_tile() +
  scale_fill_gradient2(low = "white", mid = "blue", high = "red", limits = c(0.5, 1.0)) +
  labs(title = "Voting Similarity of Countries within the Warsaw Pact", x = "", y = "") +
  theme_minimal() +
  theme(axis.text.x  = element_text(angle=90, vjust=0.5, size=8,hjust=1))
```

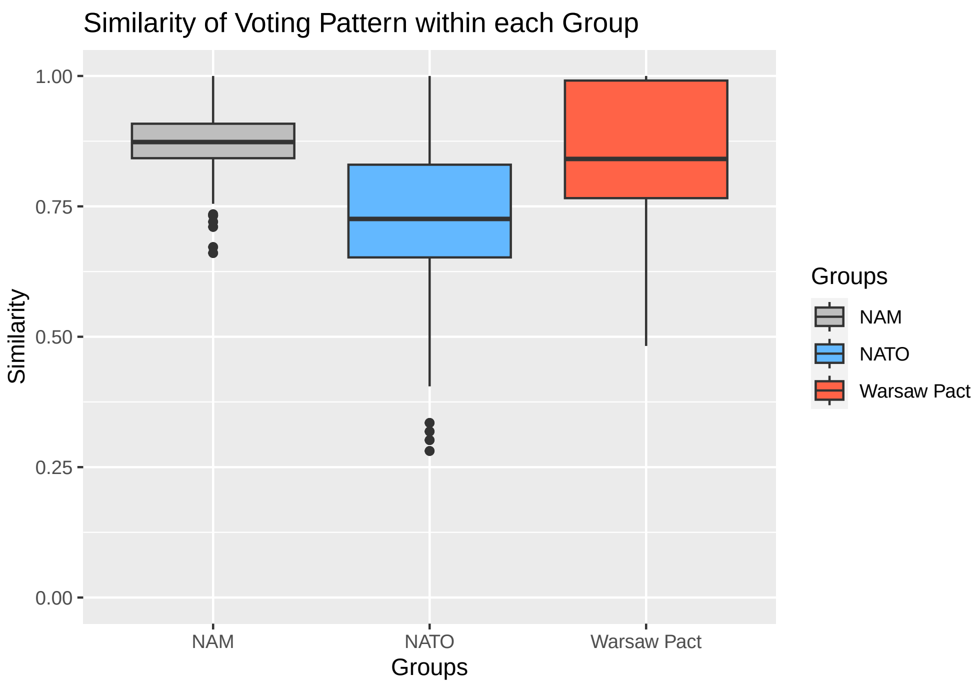



```
Countries_between_NATO_WP <- c("United States", "Russia", "Albania", "German Democratic Republic", "Fede

Countries_between_NATO_WP_tidy_matrix <- similarity_matrix[Countries_between_NATO_WP, Countries_between_
  as_tibble(rownames = "country1") %>%
  pivot_longer(-country1, names_to = "country2", values_to = "similarity")

ggplot(Countries_between_NATO_WP_tidy_matrix, aes(x = country1, y = country2, fill = similarity)) +
  geom_tile() +
  scale_fill_gradient(low = "red", high = "blue") +
  labs(title = "Voting Similarity of Countries between NATO and Warsaw Pact", x = "", y = "") +
  theme_minimal() +
  theme(axis.text.x  = element_text(angle=90, vjust=0.5, size=8,hjust=1))
```



Plot 2

```
NATO_data_frame <- as.data.frame(NATO_tidy_matrix)
WP_data_frame <- as.data.frame(WP_tidy_matrix)

nam_members <- c("Afghanistan", "Algeria", "Bangladesh", "Cuba", "Egypt", "India", "Indonesia", "Iraq",
nam_tidy_matrix <- similarity_matrix[nam_members, nam_members] %>%
  as_tibble(rownames = "country1") %>%
  pivot_longer(-country1, names_to = "country2", values_to = "similarity")
nam_data_frame <- as.data.frame(nam_tidy_matrix)

WP_data_frame_group <- WP_data_frame %>%
  mutate(Groups = "Warsaw Pact")
nam_group <- nam_data_frame %>%
  mutate(Groups = "NAM")

df_combined <- NATO_data_frame %>%
  mutate(Groups = "NATO") %>%
  rbind(WP_data_frame_group,nam_group) %>%
  group_by(Groups)

ggplot(df_combined, aes(x=Groups, y=similarity, fill=Groups)) +
  geom_boxplot() +
  ylim(0,1) +
  scale_fill_manual(values=c("gray", "steelblue1", "tomato1")) +
  labs(title="Similarity of Voting Pattern within each Group",x="Groups",y="Similarity")
```



Discussions

For the first two plots, greyed out squares indicate that the similarity of two countries are below 0.5 and considered insignificant. Blue represents high similarity with NATO states while red is high similarity within Warsaw Pact states. For NATO, the visual appearance of the plot is mostly blue, indicating that member states tend to vote in a synchronised manner. Surprisingly, the United States has the most number of greyed out squares in its row, suggesting that despite being the leader of the NATO bloc, it tended to vote dissimilarly from the NATO members. This may be because NATO states were not expected to be in sync on non-security related issues and voting difference on these issues “diluted” the degree of similarity between the US and the other states.For the plot of voting similarity within the Warsaw Pact, it also shows most countries have high voting coincidence except Albania. This may be because Albania left the pact in 1968 and was thus no longer obligated to vote in a similar way as the other pact states. For the last plot, we zoomed in on the outlier, Albania and both East and West Germany. Besides West Germany, the US has low similarity with the other Warsaw Pact states, as expected. Russia similarly only has lower similarity with the NATO states. Albania, despite having lower similarity within the Warsaw Pact compared to the other member states, still has lower similarity with the NATO states. This suggests that the strategic interests of Albania still aligned more with the pact states, regardless of the existence of the pact.

From the box plot, the median of the blocs are ordered from largest to smallest as: NAM, Warsaw Pact, NATO. In line with the findings of plot 1, NATO appeared to be voting less in-sync compared to the Warsaw Pact. And despite its name, NAM was most highly aligned in terms of its voting patterns. This may be because in the interest of neutrality when voting on issues polarised between NATO and the Warsaw Pact, NAM members may have consistently abstained and thus resulted high voting coincidence. This in combination with how both NATO has outliers where a few countries that had a strikingly opposing voting pattern compared to the rest also explains the relatively narrower similarity range of NAM compared to the two voting blocs. Although NAM has outliers as well, this did not seem to affect much of the voting pattern within NAM, given how it still has the highest median and smallest inter-quartile range. Thus, we are able to conclude that there are signs of voting coincidence within all three blocs, due to their similarity ranges being closer to 1 rather than edging towards 0.

References

1. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2111149
2. https://www.state.gov/wp-content/uploads/2022/11/Report-Voting-Practices-in-the-United-Nations-2021.pdf

Question 2: Main issues that elicit divisive or polarised voting patterns and their potential reasons

Introduction

Our second question to address would be "Are there any issues or topics that tend to elicit more divisive or polarised voting patterns, and if so, what are some potential reasons for this?" To answer this question we will be using the "importantvote", "date" and "short" entries from the roll_calls data set, the vote entries from the un_votes data set and the issue entries from the issues data set. We aim to identify the primary issues that cause polarised or divisive voting patterns in UN votes. This inquiry is crucial as it enables the recognition of underlying causes of conflicts among member states, enlightens policy-making, assists in preventing and resolving conflicts, and fosters international cooperation and collaboration. Comprehending these issues is crucial for accomplishing worldwide harmony, safety, and sustainable growth.

Methodology

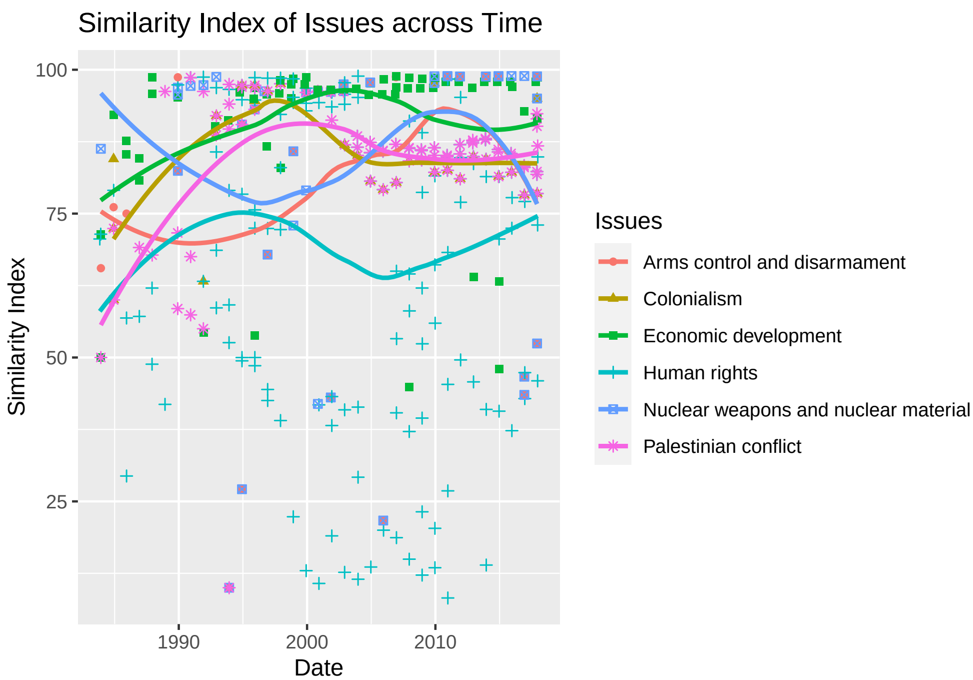
In order to measure the polarity of votes, we came up with a similarity index, which measures the extent of similarity of the votes casted within each issue. The index ranges from 0 to 1 with 0 representing very polarised votes and 1 being all votes casted to be the same. The similarity index is derived from the percentage of the absolute difference between the number of yes and no votes casted, out of the total number of effective votes casted. Effective votes is defined as the total number of yes and no votes so "abstain" votes were not included. In the first graph, we plot a scatter plot of the similarity index of each issue across the years. It is used to capture the general spread of the similarity indexes across the years. A loess regression line is then fitted for each issue, and aims to capture the time period range where each issue's similarity index begins to change. The lower the graph on the y axis, the more polarised the votes are, allowing us to identify the issues with more polarising votes.

Then, we wanted to further investigate on human rights issues. We utilised the short descriptions from the roll_call data set to identify the causes within the broader category of human rights issues. A bar plot of the top 10 human rights issues with the lowest similarity index is plotted over time. This plot helps us to identify the specific human rights issues that have more polarised votes. The specific issues that appear more frequently or have the lowest similarity index are the main causes for the polarising votes. A bar plot is a good fit for this visualisation as we can compare the extents of similarity indexes throughout the years easily since the bars are side by side. It is also suitable for us to easily extract data from the plot.

Visualisation

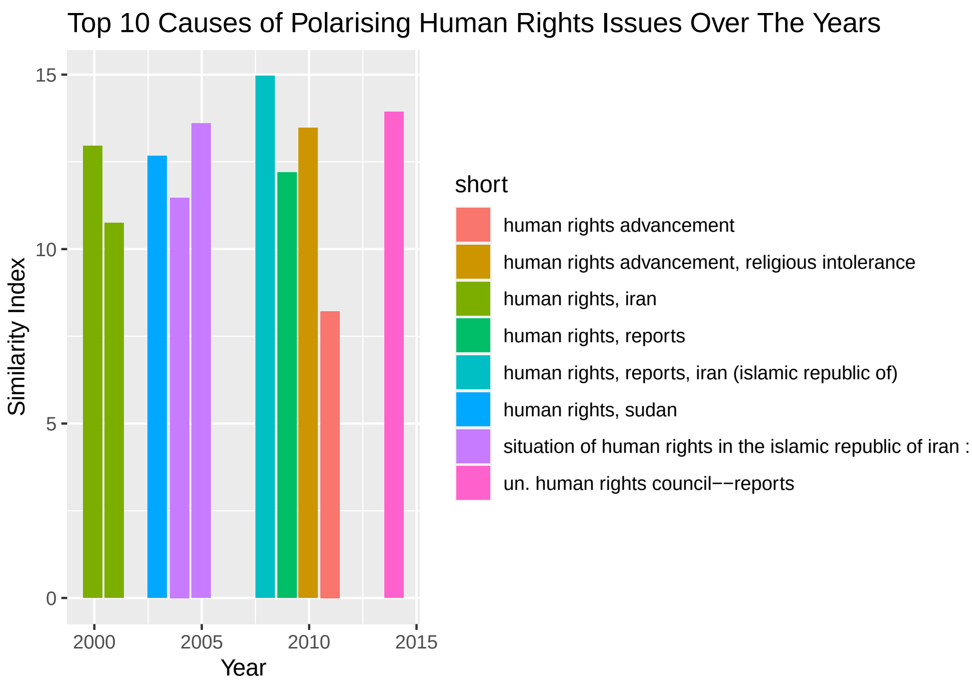
Plot 1

```r
ggplot(data = important,aes(x=date,y=similarity_index, group=issue)) +
    # facet_wrap(~issue) +
  geom_point(aes(shape=issue,color=issue),position="jitter") +
  geom_smooth(aes(color=issue),method=loess,se=FALSE) +
  labs(x="Date",y="Similarity Index",color="Issues",shape="Issues",title="Similarity Index of Issues ac
```



Plot 2

```
# Filter out issues with more polarized data (similarity_index < 50) and find the issue with the highes
short <- important %>%
  ungroup() %>%
  select(rcid,issue,date,similarity_index) %>%
  left_join(combined %>% select(rcid,short),by="rcid") %>%
  distinct() %>%
  na.omit() %>%
  filter(similarity_index < 50)

# Filter out the top 10 short descriptions with the lowest similarity index
hr <- short %>%
  filter(issue == "Human rights") %>%
  arrange(similarity_index) %>%
  top_n(-10, similarity_index) %>%
  mutate(short = tolower(short))

# Plot the bar graph of the top 10 human rights issues
ggplot(data = hr, aes(x = date, y = similarity_index, fill = short)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(x = "Year", y = "Similarity Index", title = "Top 10 Causes of Polarising Human Rights Issues Ove
```



Discussions

The first graph presented demonstrates that most issues are not highly polarised, with a similarity index of 75 or higher since 2000, except for human rights issues as seen from the consistently low similarity indexes associated with its regression line. Furthermore, the scatter plot shows that there is an observable surge of data points that fall below the regression line, and even below the 50 mark after around 2000. Most of these surges after 2000 are associated with human rights issues. When we filtered the top 10 human rights issues and plotted them on a bar graph, the results revealed the top issues occurred from 2000 to 2014, supporting our initial observations from the first graph.

According to the bar graph, the issue of human rights advancement in 2011 was found to be the most polarised with lowest similarity index. In 2011, UN's Human Rights Chief, Navi Pillay, highlighted several global challenges to human rights, including economic, climate, energy, and food crises, armed conflicts, racism, and poverty. On the positive side, civil society was mobilised to challenge oppressive governance structures in the Middle East and North Africa. However, in some regions, legitimate demands for freedom and justice were met with violence and repression.

References

1. https://www.ohchr.org/en/stories/2012/03/2011-was-year-critical-human-rights