NTNU
Kunnskap for en bedre verden

The DEMOCRAT project

Democracy Protest Actors Data

Codebook

PI:
Thea Johansen
Contact: thea.johansen@ntnu.no

10th May 2023

Table of Contents

List of Figures

List of Tables

1 Introduction

This is the codebook for the DEMOCRAT project. The purpose of the project is to map actors that participate in acts of protest and their goals in democracies that experience democratic backsliding. The efforts of this data collection will result in a dataset on actors that seek to improve or protect democracy and actors that support or push for democratic backsliding. The DEMOCRAT data can be used to identify who mobilizes for and against democracy, under what conditions they mobilize, if and under what conditions actors are more or less likely to cooperate to protect democracy, and under what conditions pro-democracy mobilization is more likely to hinder or halt democratic backsliding.

In order to collect information on protest actors and their goals in countries experiencing democratic backsliding, the DEMOCRAT projects combines information from the Episodes of Regime Transformation dataset (v2) to identify democracies experiencing episodes of democratic backsliding (Edgell et al., 2020) with event-coded data from the Integrated Crisis Early Warning System (ICEWS) database (Boschee et al., 2023). Your job as a coder is to perform systematic searches for each event in your assigned case and download newswires that can contain information on who participated in these events and what their goals were. For each event-day, you will perform one systematic search in a database for newswires called FACTIVA, and then use the newswires to code information on protest actors and their goals in R. Below is a step-for-step description of each stage of the data collection.

2 Setting up your computer

To work on DEMOCRAT, you need four things installed on your computer: Box, VPN, and R+RStudio.

2.1 Box

Firstly, you need a Box account. Box is a cloud service that meets Norwegian requirements for safe storage of the type of data we will collect in this project. Click here for instructions on how to set up Box on your computer using your NTNU license.

2.2 VPN

Secondly, in order to use FACTIVA to download newswires you will have to be connected to NTNU's network. If you want to work off campus you will need a VPN on your computer that connects you to the NTNU network. If you work on a private computer you will have to download the VPN while at campus. Click here for instructions on how to set up VPN on your computer.

2.3 R and RStudio

To code information on actors and their goals, we will use R which is a free software environment for statistical computing and graphics, and RStudio which is an integrated development environment that makes it easier to work with R. After downloading, you will only be working in RStudio.

2.3.1 Setting up R and RStudio

Download the newest version of R from the CRAN website. Press this link, choose the newest version and follow instructions for installation. Note that if you have an NTNU computer, you will have to go through some additional steps before you can do this which is described in the section

below. After you have downloaded R, it is time to download RStudio. RStudio is what you will be coding in. Click on this link and download the newest free version of RStudio. Once you have downloaded both you are ready to get started on the data collection!

2.3.2 If you have a NTNU computer

All NTNU computers come without full admin rights when we receive them from IT, and unfortunately this can make certain functions and packages in R very slow or unable to run.

1. Make sure you have administration rights for your NTNU laptop (full admin rights)! If you don't have it, you can get it by applying for it using the form on NTNU Hjelp (this might take a couple of hours). Do not do anything else before you get this, follow the instructions from IT for when to restart your computer.

2. Go to the control panel on your laptop, search environment variables, click edit and add new variable called R_LIBS_USER, and give it a value path to where you want to store R stuff, for instance "C:/Users/yourusername/R" (make a folder first if you don't have one you want to use). Restart your computer after this.

3. Download R and RStudio as described in the previous section. Just doublecheck before downloading that this is now stored on your server **C** and not on your network disk **M**.

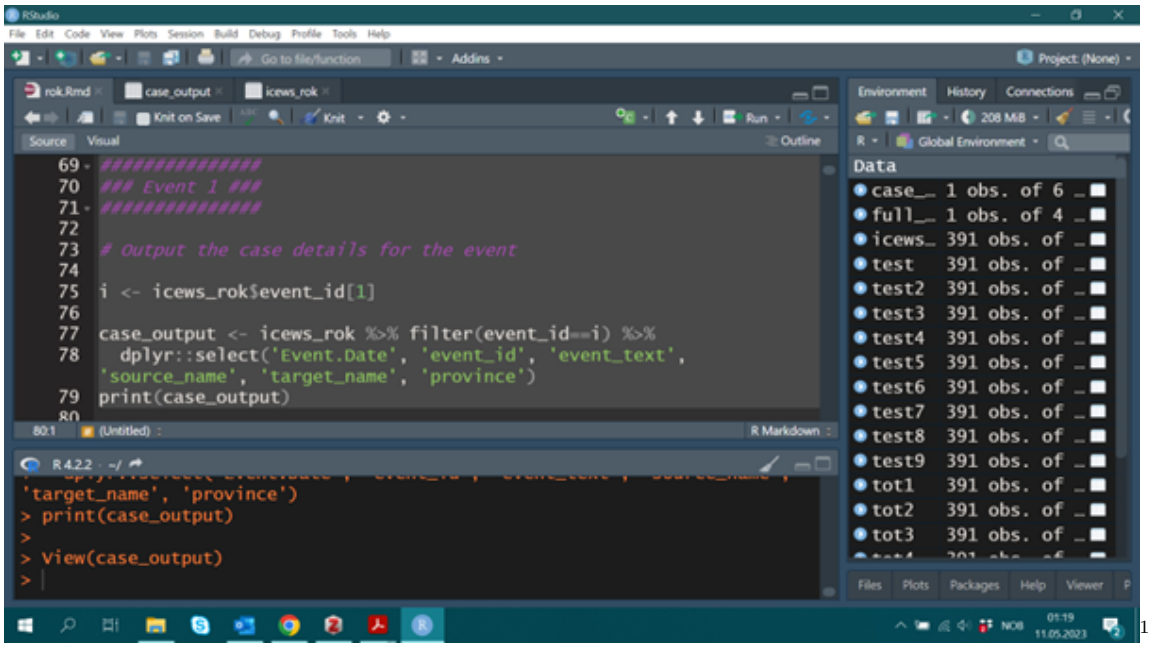
4. Open RStudio and type in the console .libPaths() and hit enter. Your reply should look something like this: [1] "C:/Users/yourusername/R" "C:/Program Files/R/R-4.1.0/library"

3 Getting started

You will work on one country-case at the time. In the DEMOCRAT Box you will find that your country has a RMarkdown file, and a country folder which contains one folder called events and one folder called newswires.

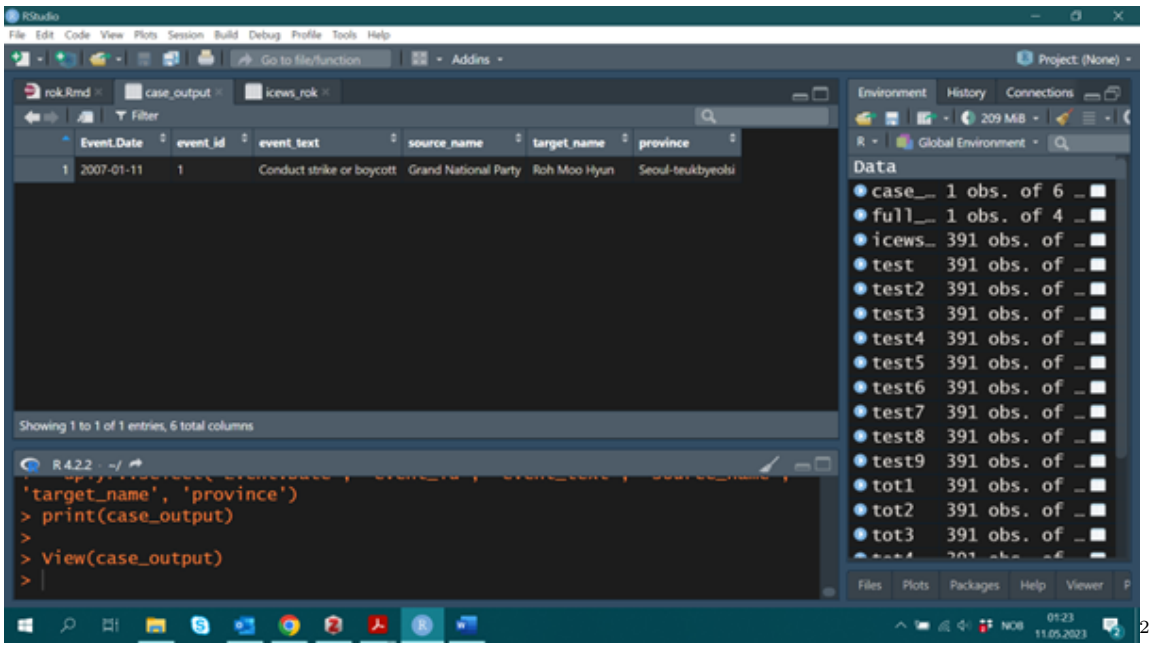
4 Country RmD-file

When you are in R, there is a bunch of code at the start of the Markdown-file. Run everything including code line 79 as illustrated in the picture below:



[1]

This prints the information you need to specify systematic newswires search for event 1. Note the date information for the case output which you will use to specify your search in your next step in FACTIVA.



[2]

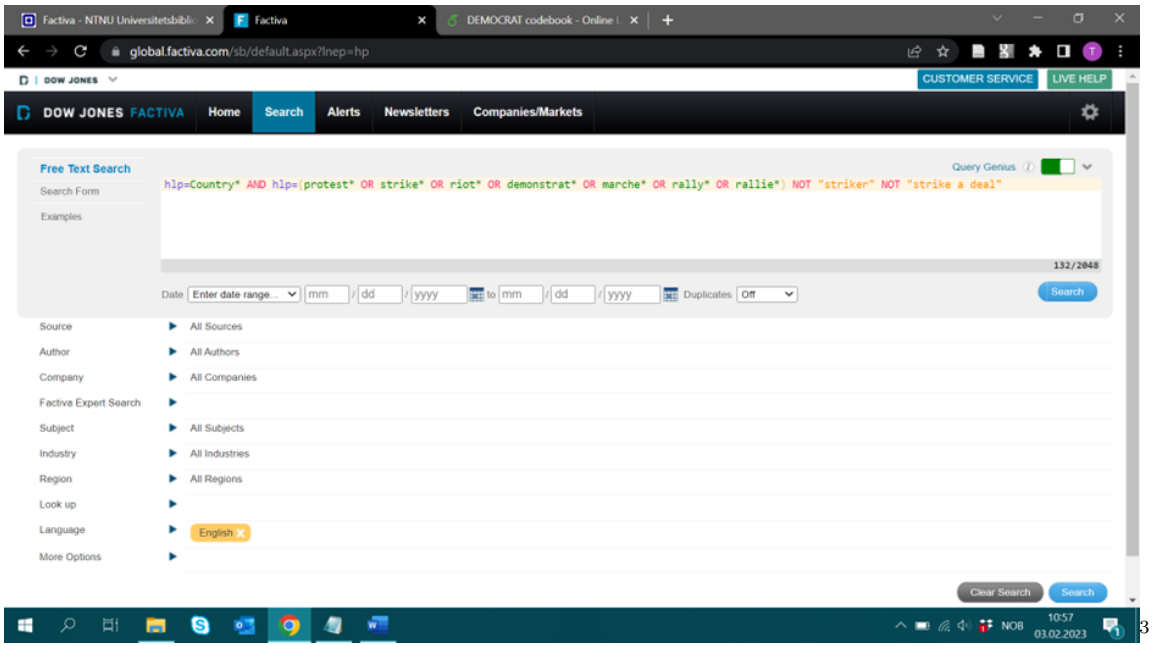
5 FACTIVA

FACTIVA is the news database you will perform systematic searches for. Click here for a link to log into FACTIVA. Note that you can only use FACTIVA when connected to the NTNU network, so make sure to use VPN if you are working off campus.

The FACTIVA interface looks like the picture below. For each event-day, you will enter the same search string: **hlp=Country* AND hlp=(protest* OR strike* OR riot* OR demonstrat* OR marche* OR rally* OR rallie*) NOT "striker" NOT "strike a deal" AND**

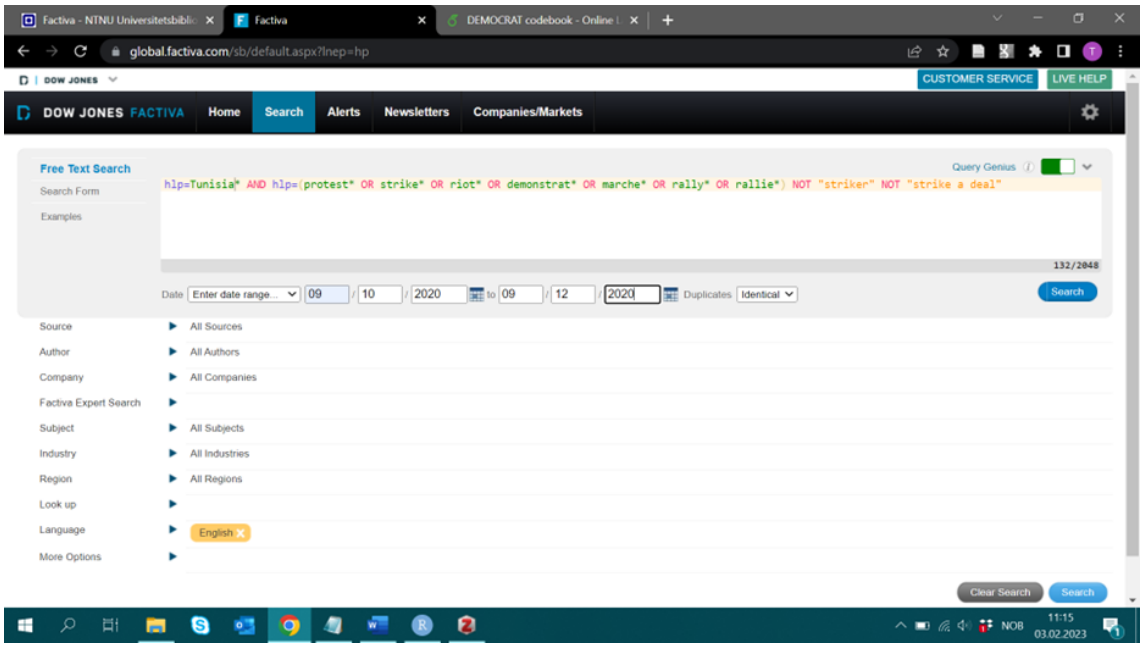
[1] Pictures will be updated
[2] Pictures will be updated

(SC=AFPR OR SC=APRS OR SN="BBC" OR SC=LBA). In addition, we will specify to search for newswires from six region-specific sources in the search string which will vary for each case. You can find the search string for each country in Box. You should then make two changes to the search string. First, change **Country*** to the name of the case that you are coding. Note that the asterisk is included because we want to be able to pick up both the name of the country and its demonym(s). This means that if the name of the country is Tunisia and Tunisia's demonym is Tunisian, we can write **Tunisia***. However, if the name of the country is Norway and the demonym is Norwegian, you should write **Norw***.



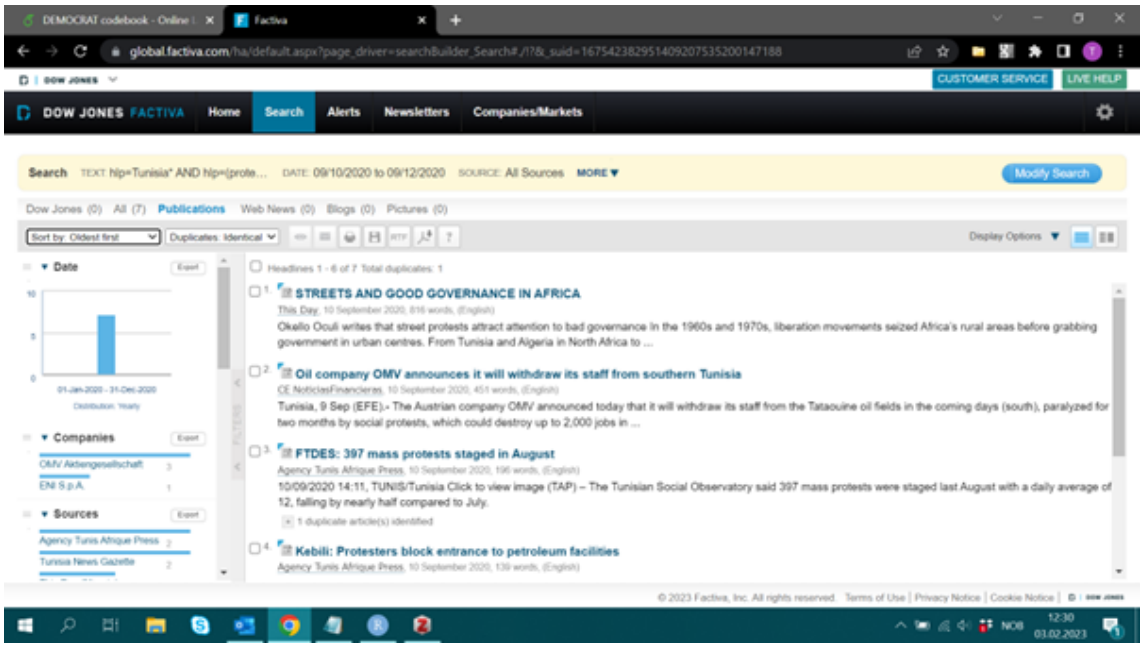
[3]

The next step is to specify the dates for the search. Information on date for each event-day is found in the case_output for each event. In FACTIVA, press enter date range as in the picture below and enter the day the event starts as the starting date and the day after the event as the end date. Duplicates should be set to identical as in the example below. Then, you can just hit search!

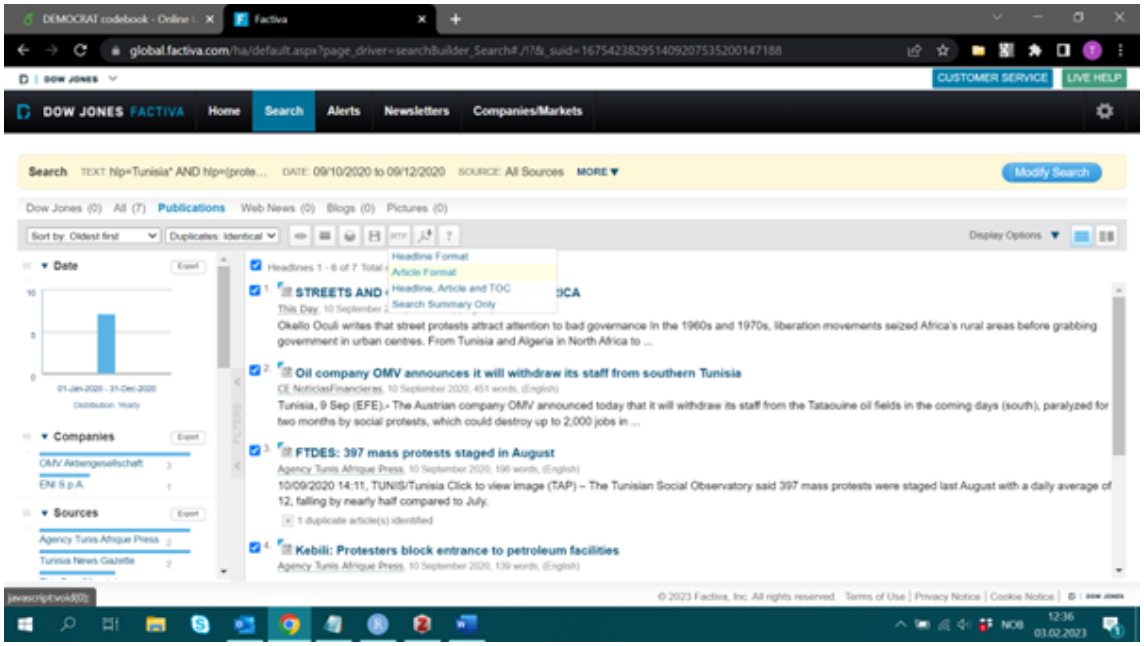


[3]Pictures will be updated

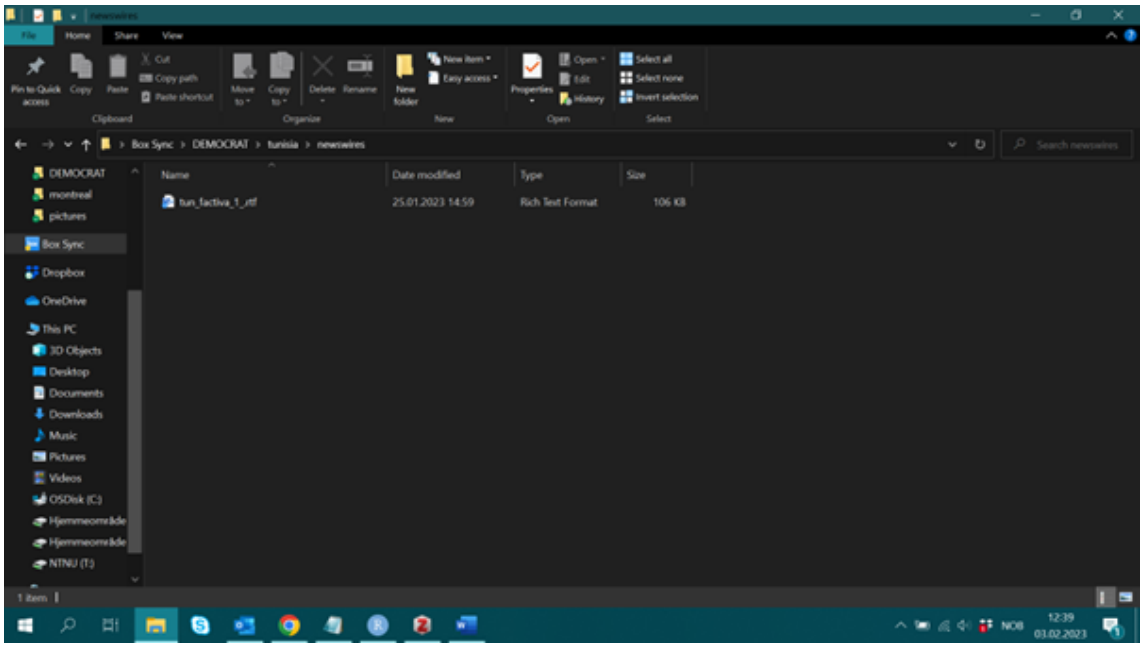
Next, it is time to download the newswires as a rich text file (RTF) which you will later import in R to code actors and goals. First, you need to arrange the reports by date by clicking **Sort by: Oldest first** as illustrated in the picture below.



Once you have sorted the newswires, you click on **Headlines** to mark all newswires. Then you press the **RTF** icon and choose article format in the drop down menu. If there are more than a 100 newswires, we will only download the first 100.



Save the RTF in the newswires folder in Box for the countrycase you are working on. Note that the file should always be named in this format: **xxx_factiva_y_.rtf** where xxx represents the three letter countrycode and y is the **event_id** of the protest-day. For example, for the first protest-day for Tunisia the saved file should be named and stored like this:



Excellent! You are now ready to return to R and begin coding your first case.

6 Returning to R

7 Variables

When you have downloaded the newswires from Factiva, it is time to upload it in R by running this line of code:

After that, we take a look at the case_output for the event. Using only the information provided by the news reports, we will code three set of variables.

7.1 Democracy protest

Binary variable 0-1 where 1 indicates that the protest is related to the protection or improvement of democratic norms, rights, and/or liberties. While we code specific information on this on the actor level in the next step, we still want a binary variable we can use to sort whether a protest can be considered a pro-democracy protest regardless of actor participation.

7.2 Actor variable

The actor variable is a variable coded on a sentence level, and consists of three components: a) a unique organization code + b) an indicator for the goal of the protesters + c) A binary measure for whether the demand was maximalist or not. These three components are combined like this: ORG-000_00. Below is a description of how to assign the organization-id before the underscore, and how to assign the two digits after the underscore which codes information on the protest goal.

7.2.1 ORG-ID

An unique organization code for each identified formal organization. The variable Consists of the three letter countrycode which represents the country where the organization was founded.

Assign "TUN" for a Tunisian organization, "ROK" for South Korean, "POL" for Polish, "ECU" for Ecuadorian, "BRA" for Brazilian, and "HUN" for Hungarian. If the organization's country-of-origin is different from the case you have been assigned, you should use the countrycode for the country-of-origin; for example, an Egyptian organization in Tunisia should be coded as EGY-001.

After the three-letter countrycode, you add a dash and then assign a unique three digit organization code. For example, TUN-001 indicates that the organization is a Tunisian organization, and the 001 is a unique id for that organization. The next organization you find should be given a new id, for instance TUN-002.

7.2.2 Protest goal

The next step of the actor variable is to add the organization's protest goal for that specific event. This means that depending on the event, ORG-001 can be assigned a different goal for each event depending on the goal. We will add information on whether the goal was unrelated to democracy *and* whether the goal can be considered maximalist or not. Below are three tables that explain the to sets of goal types and how to combine them.

Table 1: Variable: Protest goal

Category	Label	Description	Examples
0	Unrelated to democracy	These are demands that are not directly related to the protection or improvement of democratic norms, rights, and/or liberties	The General Labour Union called for a general strike for better pay.
1	Pro-democracy	Demands or expressed support for action that if implemented would protect or improve democratic norms, rights, and/or liberties	The General Labour Union demanded that the police respected their constitutional right to organize strikes. The General Labour Union called for a new law that would protect the independence of labour organizations
2	Pro-regime	Statements and slogans that express support for the government, or that call for increased power to the executive branch of government	The General Labour Union expressed support for the government's decision to crack down on the counter-protest.

7.2.3 Maximalist

Table 2: Variable: Maximalist

Category	Label	Description
1	Maximalist	Demands that if implemented would change who has access to executive power
0	Non-maximalist	Demands that if implemented would not directly impact who has access to executive power

7.2.4 Combining protest goal with maximalist

Table 3: Possible goal description combination

	Non-maximalist: 0	Maximalist: 1
Unrelated: 0	00	01
Pro-democratic: 1	10	11
Pro-regime: 2	20	21

7.2.5 Example on how to combine org-id and protest goals

First event: identify org

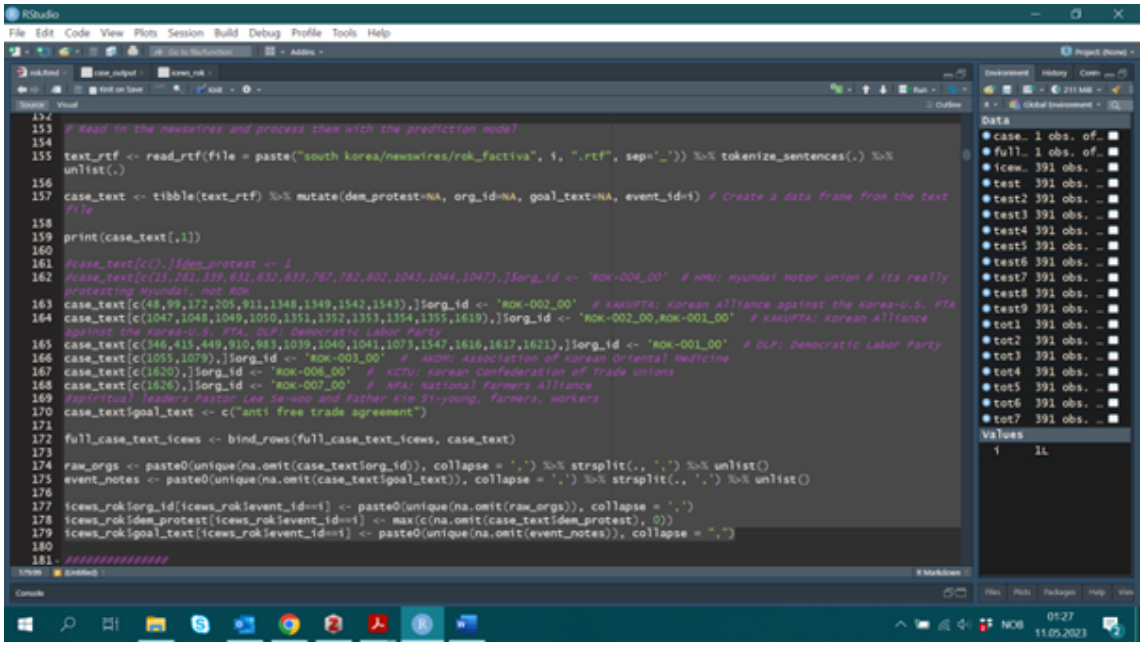
Second event: same org, different goal

Third event: how to combine multiple actors

7.3 Goal text

The goal text variable is a string variable called goal_text where we will in as few words as possible summarize the main goal of the protesters. For instance, if the protest is about protecting abortion rights, we can add case_text$goal_text ¡- "pro abortion". If the event is about higher pay and economic conditions, we can enter case_text$goal_text ¡- c("economic issues").

7.4 Example of a coded event



7.5 Adding variable names

We also want to assign a name to each organization. To do this, we scroll to the end of the RMarkdown-file to this snippet of code and add an abbreviation and the actual name for each org-id like this:

```
1783 #
1784 # # Add Organization Names
1785 #
1786 rok_final_list$org.name <- NA
1787 #
1788 rok_final_list$org.name[rok_final_list$org_id=='ROK-001'] <- "DLP: Democratic Labor
     Party"
1789 rok_final_list$org.name[rok_final_list$org_id=='ROK-002'] <- "KAKUFTA: Korean Alliance
     against the Korea US FTA"
1790 rok_final_list$org.name[rok_final_list$org_id=='ROK-003'] <- "AKOM: Association of
     Korean Oriental Medicine"
1791 rok_final_list$org.name[rok_final_list$org_id=='ROK-004'] <- "HMU: Hyundai Motor
     Union"
1792
```

7.6 Adding variable types

Finally, we also want to separate between different types of organizations. To do this, we have to assign a value to the variable organization type. We do this as illustrated below:

```
1796  #
1797  # # Initial encoding of organization type
1798  #
1799  rok_final_list$org_type <- NA
1800  #
1801  rok_final_list$org_type[rok_final_list$org_id=='ROK-001'] <- 1
1802  rok_final_list$org_type[rok_final_list$org_id=='ROK-002'] <- 8
1803  rok_final_list$org_type[rok_final_list$org_id=='ROK-003'] <- 2
1804  rok_final_list$org_type[rok_final_list$org_id=='ROK-004'] <- 2
```

Here is a table of the different organization types you can assign. We determine the organization type by assessing what the main goal of the organization is.

Table 4: Organization type

Code	Type	Org. Goal	Example
1	Political Party	Win elections	
2	Labour Union	Maintain or improve conditions of employment	
3	Women's	Women's rights	
4	LHBT+	LHBT+ rights	
5	Student or Youth	Students/youth rights	
6	Religious	Religious goals	
7	Revolutionary	Regime change	
8	Other CSOs	Other	

8 Wrapping up

Bibliography

Boschee, E., Lautenschlager, J., O'Brien, S., Shellman, S., Starz, J., & Ward, M. (2023). ICEWS Coded Event Data [Type: dataset]. https://doi.org/10.7910/DVN/28075

Edgell, A. B., Maerz, S. F., Maxwell, L., Morgan, R., Medzihorsky, J., Wilson, M. C., Boese, V. A., Hellmeier, S., Lachapelle, J., Lührmann, A., & Lindberg, S. I. (2020). Episodes of Regime Transformation Dataset (v2) Codebook. https://www.v-dem.net/documents/9/ert_codebook.pdf

Appendix

A Other questions

What if I have more than 100 newswires in FACTIVA?
If you have 100 newswires for a protest-day, we would first try to limit the newswires by excluding irrelevant newswires. By taking a quick glance at the headlines you can identify if there are some type of reports which are clearly unrelated to the event. If so, we can add some additional terms to the search string in FACTIVA to further specify this. For example, if there are several newswires on new research that demonstrate the effect of a new drug called CURE, you can add **NOT "CURE"** to the end of the search string and search again. If there are still 100 newswires, we will limit the search to only the day of the event. If there are still 100 newswires, we will only code the first 50 and the final 50. You do this by downloading the first 50 in one rtf-file and the final 50 in another rtf-file, and then copy-paste the content of the last one into the first rtf. Save, then code as normal.