

# Climate Negotiations Poster Data

Bronwen Boyd

## Webscraping the Sabin Center for Climate Change Law Database

```
#Accessing the necessary packages
library(rvest)
library(xml2)
library(purrr)
library(dplyr)
```

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

filter, lag

The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

```
library(tidyverse)
```

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v forcats   1.0.0      v stringr   1.5.1
v ggplot2   3.5.1      v tibble    3.2.1
v lubridate 1.9.3      v tidyr     1.3.1
v readr     2.1.5
```

```
-- Conflicts ----- tidyverse_conflicts() --
x dplyr::filter()      masks stats::filter()
x readr::guess_encoding() masks rvest::guess_encoding()
x dplyr::lag()          masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

```
library(stringr)
library(tinytex)
```

Warning: package 'tinytex' was built under R version 4.4.1

```
#retrieving the main webpage for Sabin Center for Climate Change Law Database
main_webpage_1 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-foreign-countries")

#retrieving the first sheet of subpages on this main webpage. Using Cmd + option + i to open the developer tools

subpage_links <- main_webpage_1 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

#This worked! I can now type the code for the remaining 34 sheets of sublinks for the individual cases

main_webpage_2 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-foreign-countries")
subpage_links_2 <- main_webpage_2 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_3 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-foreign-countries")
subpage_links_3 <- main_webpage_3 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_4 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-foreign-countries")
subpage_links_4 <- main_webpage_4 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_5 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-foreign-countries")
subpage_links_5 <- main_webpage_5 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")
```

```

main_webpage_6 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-
subpage_links_6 <- main_webpage_6 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_7 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-
subpage_links_7 <- main_webpage_7 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_8 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-
subpage_links_8 <- main_webpage_8 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_9 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-
subpage_links_9 <- main_webpage_9 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_10 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-
subpage_links_10 <- main_webpage_10 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_11 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-
subpage_links_11 <- main_webpage_11 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_12 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-
subpage_links_12 <- main_webpage_12 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_13 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-
subpage_links_13 <- main_webpage_13 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_14 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-
subpage_links_14 <- main_webpage_14 %>%

```

```

html_nodes("a.read-more") %>%
html_attr("href")

main_webpage_15 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_15 <- main_webpage_15 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_16 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_16 <- main_webpage_16 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_17 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_17 <- main_webpage_17 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_18 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_18 <- main_webpage_18 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_19 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_19 <- main_webpage_19 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_20 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_20 <- main_webpage_20 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_21 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_21 <- main_webpage_21 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_22 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_22 <- main_webpage_22 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

```

```

main_webpage_23 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_23 <- main_webpage_23 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_24 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_24 <- main_webpage_24 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_25 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_25 <- main_webpage_25 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_26 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_26 <- main_webpage_26 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_27 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_27 <- main_webpage_27 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_28 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_28 <- main_webpage_28 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_29 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_29 <- main_webpage_29 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_30 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_30 <- main_webpage_30 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_31 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_31 <- main_webpage_31 %>%

```

```

html_nodes("a.read-more") %>%
html_attr("href")

main_webpage_32 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_32 <- main_webpage_32 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_33 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_33 <- main_webpage_33 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_34 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_34 <- main_webpage_34 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

main_webpage_35 <- read_html("https://climatecasechart.com/non-us-case-category/suits-against-us")
subpage_links_35 <- main_webpage_35 %>%
  html_nodes("a.read-more") %>%
  html_attr("href")

#combining all 35 web sheet subpage URLs into one character vector using direct concatenation
all_subpage_links <- c(subpage_links, subpage_links_2,subpage_links_3, subpage_links_4, subpage_links_5, subpage_links_6, subpage_links_7, subpage_links_8, subpage_links_9, subpage_links_10, subpage_links_11, subpage_links_12, subpage_links_13, subpage_links_14, subpage_links_15, subpage_links_16, subpage_links_17, subpage_links_18, subpage_links_19, subpage_links_20, subpage_links_21, subpage_links_22, subpage_links_23, subpage_links_24, subpage_links_25, subpage_links_26, subpage_links_27, subpage_links_28, subpage_links_29, subpage_links_30, subpage_links_31, subpage_links_32, subpage_links_33, subpage_links_34, subpage_links_35)

#Now going into these compiled URLs and retrieving the information I need from each of the
scrape_case_info <- function(all_subpage_links) {
  webpage <- read_html(all_subpage_links)
  case_title <- webpage %>% html_node("h1") %>% html_text(trim = TRUE)
  filing_date <- webpage %>% html_node("span.highlight") %>% html_text(trim = TRUE)
  jurisdiction1 <- webpage %>% html_nodes(".entry-taxonomy-item a") %>% html_text(trim = TRUE)
  # .[grepl("jurisdiction",.)]

  # If jurisdiction is empty, replace it with NA
  if (length(jurisdiction1) == 0) {
    jurisdiction1 <- NA
  }
}

#Code for extracting information on case status

```