Data Analysis Report

Michael V Cumbo

April 13, 2024

1 Introduction

This paper was written in response to the United Auto Workers strike and the SAG-AFTRA strike of 2023. The goal of gathering the data is to contextualize the state of trade union power in the United States, blending data analysis with a literature review. The data in this paper will contextualize union power in a select number of nation-states, adding perspective to the modes of influence unions have.

This report is a month to month analysis of labor action tracker newly recorded data. The report should give an assessment of trends in the strike activity. The regions in focus will be the DMV which consist of The District of Columbia, Maryland and Virginia.

2 Data Prep

```
source("~/Lab2/Trade_Union_Global_Analysis/summary_analysis/Enviroment_Setup.R")

suppressPackageStartupMessages({
  library(ggplot2)
  library(dplyr)
})

# connect to the database and pull out the sheet
setwd("~/workbook")
con <- dbConnect(RSQLite::SQLite(), "trade_union_data.db")
dbListTables(con)


##  [1] "CBCR"
##  [2] "CollectiveBargaining"
##  [3] "ILRLaborActionTracker"
##  [4] "IRL&ScrappedDataComparision"
##  [5] "LAT-02.19.24"
##  [6] "LAT-03.04.24"
##  [7] "LAT-04.01.24"
##  [8] "State_Union_Coverage_Density_1977-2021"
##  [9] "State_Union_Membership_Density_1964-2021"
## [10] "Strikes_United_States"
## [11] "TUDR"
## [12] "TradeUnionDensity"
## [13] "WorkplaceRights"
## [14] "raw_strike_table"
## [15] "rawonlystrike_table"
## [16] "state_union_coverage_density_long"
## [17] "state_union_membership_density_long"
## [18] "state_uniondc_join"
```

```
## [19] "strike_table"
## [20] "summarytable_IRLvSData"


february <- as_tibble(dbReadTable(con, "LAT-02.19.24"))
dbDisconnect(con)
```
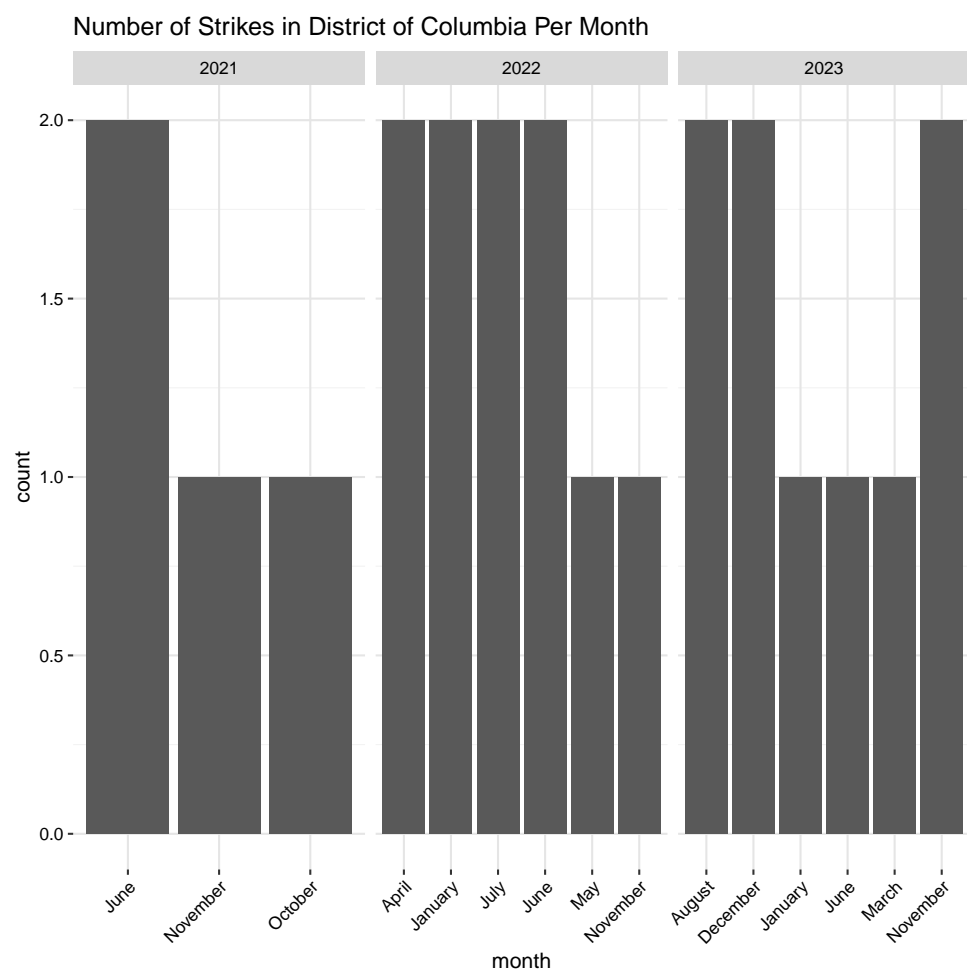
```
# Convert Timestamp column to POSIXct format
february$Timestamp <- as.POSIXct(february$Timestamp,
  format = "%m/%d/%Y %H:%M:%S"
)

# mutate zip code to a character
february <- february %>%
  mutate(
    ZipCode = as.character(ZipCode),
    BargainingUnitSize = parse_number(as.character(BargainingUnitSize)),
    ApproximateNumberofParticipants =
      parse_number(as.character(ApproximateNumberofParticipants)),
    Date = format(Timestamp, "%Y-%m-%d"),
    month = format(Timestamp, "%B"),
    Year = format(Timestamp, "%Y"),
    DurationAmount = as.integer(DurationAmount)
  )
```
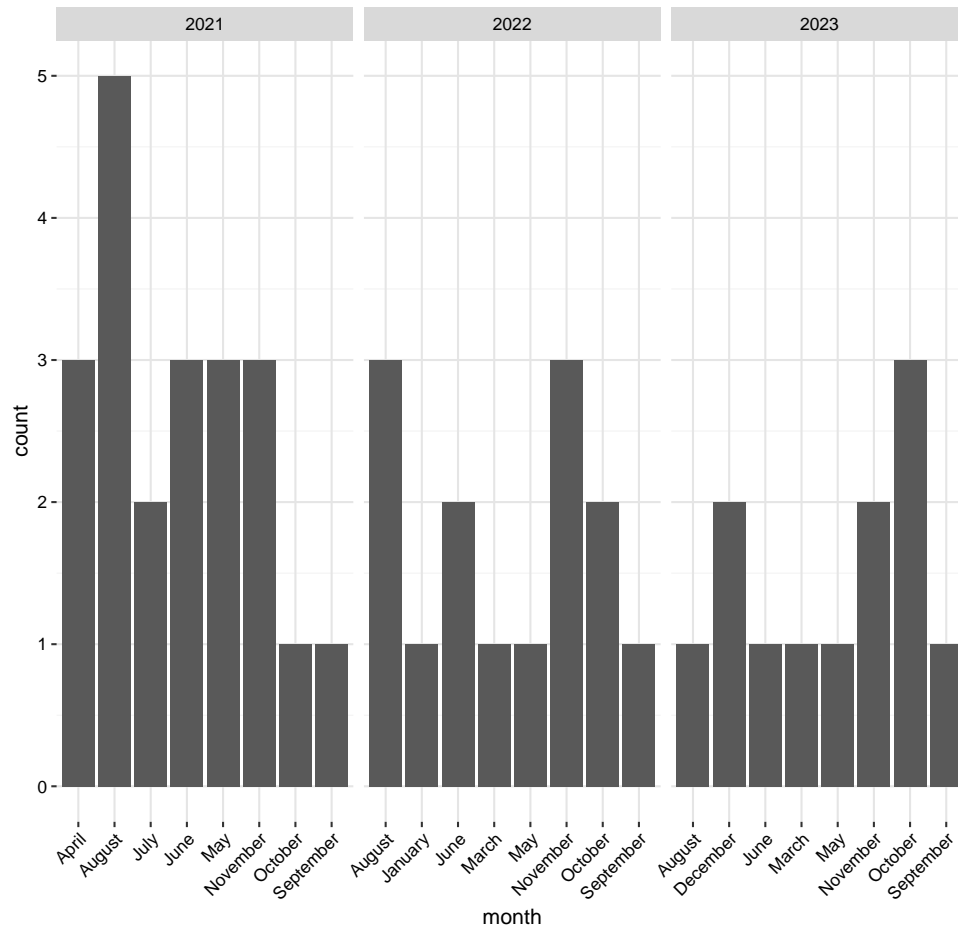
```
## [1] "the total number of available data points in the BargainingUnitSize vector is 211"
## [1] "the total number of missing/blank data points in the BargainingUnitSize vector is 2655"
## [1] "the total number of missing/blank data points in the notes vector is 380"
```

3 Data Analysis

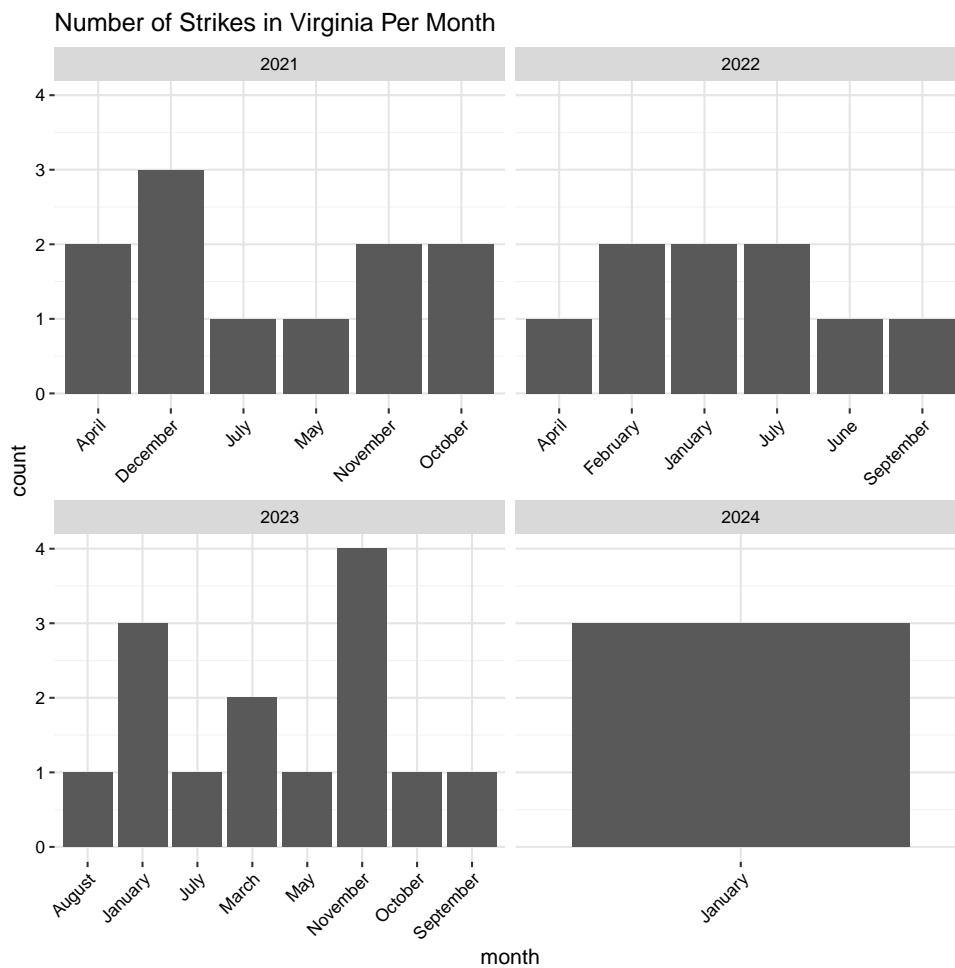Number of Strikes in District of Columbia Per Month


```
## The number of active labor organizations during the selected
##  period of 2021, 2022, 2023, 2024 in District of Columbia is 18 .
## # A tibble: 3 x 4
##   State                Year  labor_org_count strikes
##   <chr>                <chr>           <int>   <int>
## 1 District of Columbia 2021                3       4
## 2 District of Columbia 2022                7      10
## 3 District of Columbia 2023                8       9
```

Number of Strikes in Maryland Per Month


```
## The number of active labor organizations during the selected
##  period of 2021, 2022, 2023, 2024 in Maryland is 36 .
## # A tibble: 3 x 4
##   State    Year  labor_org_count strikes
##   <chr>    <chr>           <int>   <int>
## 1 Maryland 2021               11      21
## 2 Maryland 2022               14      14
## 3 Maryland 2023               11      12
```

Number of Strikes in Virginia Per Month


```
## The number of active labor organizations during the selected
##  period of 2021, 2022, 2023, 2024 in Virginia is 28 .
## # A tibble: 4 x 4
##   State    Year  labor_org_count strikes
##   <chr>    <chr>           <int>   <int>
## 1 Virginia 2021                8      11
## 2 Virginia 2022                7       9
## 3 Virginia 2023               12      14
## 4 Virginia 2024                1       3
```