

Gerrymandering Metrics

2024 Election Results and the 2024 District Map (Part 4)

```
library(sf)
```

```
Linking to GEOS 3.13.1, GDAL 3.11.0, PROJ 9.6.0; sf_use_s2() is TRUE
```

```
library(tidyverse)
```

```
— Attaching core tidyverse packages ————— tidyverse 2.0.0
—
✓ dplyr      1.1.4      ✓ readr      2.1.5
✓ forcats    1.0.1      ✓ stringr    1.5.2
✓ ggplot2    4.0.0      ✓ tibble     3.3.0
✓ lubridate  1.9.4      ✓ tidyr      1.3.1
✓ purrr      1.1.0
```

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

```
votes_precinct <- read_csv("data/g24_sov_by_g24_svprec_clean.csv")
```

```
Rows: 51123 Columns: 76
```

```
— Column specification
```

```
Delimiter: ","
```

```
chr (5): FIPS, SVPREC, SVPREC_KEY, ELECTION, GEO_TYPE
```

```
dbl (71): COUNTY, ADDIST, CDDIST, SDDIST, BEDIST, TOTREG, DEMREG, REPREG,
AI...
```

```
i Use `spec()` to retrieve the full column specification for this data.
```

```
i Specify the column types or set `show_col_types = FALSE` to quiet this
message.
```

```

votes_cong <- votes_precinct |>
  filter(ELECTION == "G24" & !is.na(CDDIST)) |>
  group_by(CDDIST) |>
  summarize(DEM = sum(CNGDEM01, na.rm = TRUE), REP = sum(CNGREP01, na.rm =
TRUE))

head(votes_cong)

```

```

# A tibble: 6 × 3
  CDDIST      DEM      REP
  <dbl>    <dbl>    <dbl>
1     0 52843210 34247705
2     1  110472   208150
3     2  272384   106407
4     3  187960   233895
5     4  227321   114644
6     5  134467   214223

```

```

votes_cong <- votes_cong |>
  mutate(DEM_share = DEM / (DEM + REP))

mean_median_diff <- mean(votes_cong$DEM_share) - median(votes_cong$DEM_share)
mean_median_diff

```

```
[1] -0.01163046
```

```

wasted_votes <- function(votesA, votesB) {
  total <- votesA + votesB
  margin <- abs(votesA - votesB)
  half <- total / 2
  ifelse(votesA > votesB, votesA - half, votesA)
}

votes_cong <- votes_cong |>
  mutate(wasted_DEM = wasted_votes(DEM, REP),
         wasted_REP = wasted_votes(REP, DEM))

efficiency_gap <- sum(votes_cong$wasted_REP - votes_cong$wasted_DEM) /
sum(votes_cong$DEM + votes_cong$REP)
efficiency_gap

```

```
[1] 0.2596723
```

```
tibble(Metric = c("Mean-Median Difference", "Efficiency Gap"),
       Value = c(mean_median_diff, efficiency_gap))
```

```
# A tibble: 2 × 2
  Metric      Value
  <chr>      <dbl>
1 Mean-Median Difference -0.0116
2 Efficiency Gap          0.260
```

2024 Election Results and the proposed 2025 District Map (Part 6)

```
sr_shp <- st_read("data/shapefiles/srprec_state_g24_v01_shp.shp") |>
  st_transform(3310) |>
  st_set_precision(1) |>
  st_make_valid() |>
  st_collection_extract("POLYGON")
```

```
Reading layer `srprec_state_g24_v01_shp' from data source
`C:\Users\jiaaa\Desktop\gerrymandering-
Selina568\data\shapefiles\srprec_state_g24_v01_shp.shp'
using driver `ESRI Shapefile'
```

```
Warning in CPL_read_ogr(dsn, layer, query, as.character(options), quiet, :
GDAL
Message 1:
C:\Users\jiaaa\Desktop\gerrymandering-
Selina568\data\shapefiles\srprec_state_g24_v01_shp.shp
contains polygon(s) with rings with invalid winding order. Autocorrecting
them,
but that shapefile should be corrected using ogr2ogr for example.
```

```
Simple feature collection with 24145 features and 6 fields
Geometry type: MULTIPOLYGON
Dimension: XY
Bounding box: xmin: -124.482 ymin: 32.52883 xmax: -114.1312 ymax: 42.0095
Geodetic CRS: NAD83
```

```
cd_ab604 <- st_read("data/shapefiles/AB604.shp") |>
  st_transform(3310)
```

```
Reading layer `AB604' from data source
`C:\Users\jiaaa\Desktop\gerrymandering-Selina568\data\shapefiles\AB604.shp'
```

```

    using driver `ESRI Shapefile'
Simple feature collection with 52 features and 15 fields
Geometry type: MULTIPOLYGON
Dimension:      XY
Bounding box:   xmin: -13857270 ymin: 3832931 xmax: -12705030 ymax: 5162404
Projected CRS: WGS 84 / Pseudo-Mercator

```

```
sr_votes <- read_csv("data/state_g24_sov_data_by_g24_srprec.csv")
```

```

Rows: 25245 Columns: 76
— Column specification

```

```

Delimiter: ","
chr (49): FIPS, SRPREC, ELECTION, SRPREC_KEY, GEO_TYPE, ASSAIP01,
ASSDEM01, ...
dbl (27): COUNTY, ADDIST, CDDIST, SDDIST, BEDIST, TOTREG, DEMREG, REPREG,
AI...

```

```

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this
message.

```

```
sr_data <- left_join(sr_shp, sr_votes, by = c("SRPREC_KEY" = "SRPREC_KEY"))
```

```
votes_precinct <- read_csv("data/g24_sov_by_g24_svprec_clean.csv")
```

```

Rows: 51123 Columns: 76
— Column specification

```

```

Delimiter: ","
chr (5): FIPS, SVPREC, SVPREC_KEY, ELECTION, GEO_TYPE
dbl (71): COUNTY, ADDIST, CDDIST, SDDIST, BEDIST, TOTREG, DEMREG, REPREG,
AI...

```

```

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this
message.

```

```

votes_cong <- votes_precinct |>
  filter(ELECTION == "G24" & !is.na(CDDIST)) |>
  group_by(CDDIST) |>
  summarize(DEM = sum(CNGDEM01, na.rm = TRUE),
            REP = sum(CNGREP01, na.rm = TRUE))

```

```
sr_data <- sr_data |>
  mutate(CNGDEM01 = as.numeric(CNGDEM01),
         CNGREP01 = as.numeric(CNGREP01))
```

Warning: There were 2 warnings in `stopifnot()`.
 The first warning was:
 i In argument: `CNGDEM01 = as.numeric(CNGDEM01)`.
 Caused by warning:
 ! NAs introduced by coercion
 i Run `dplyr::last_dplyr_warnings()` to see the 1 remaining warning.

```
reallocated_DEM <- st_interpolate_aw(sr_data["CNGDEM01"], cd_ab604, extensive
= TRUE)
```

Warning in st_interpolate_aw.sf(sr_data["CNGDEM01"], cd_ab604, extensive = TRUE): st_interpolate_aw assumes attributes are constant or uniform over areas of x

```
reallocated_REP <- st_interpolate_aw(sr_data["CNGREP01"], cd_ab604, extensive
= TRUE)
```

Warning in st_interpolate_aw.sf(sr_data["CNGREP01"], cd_ab604, extensive = TRUE): st_interpolate_aw assumes attributes are constant or uniform over areas of x

```
cd_ab604$DEM <- reallocated_DEM$CNGDEM01
cd_ab604$REP <- reallocated_REP$CNGREP01
```

```
votes_cong_ab604 <- cd_ab604 |>
  st_drop_geometry() |>
  mutate(TOTAL = DEM + REP, DEM_share = DEM / TOTAL) |>
  filter(TOTAL > 0)
```

```
mean_median_diff_ab604 <- mean(votes_cong_ab604$DEM_share, na.rm = TRUE) -
  median(votes_cong_ab604$DEM_share, na.rm = TRUE)
```

```
wasted_votes <- function(votesA, votesB) {
  total <- votesA + votesB
  half <- total / 2
  ifelse(votesA > votesB, votesA - half, votesA)
}
```

```

votes_cong_ab604 <- votes_cong_ab604 |>
  mutate(wasted_DEM = wasted_votes(DEM, REP),
         wasted_REP = wasted_votes(REP, DEM))

efficiency_gap_ab604 <- sum(votes_cong_ab604$wasted_REP -
votes_cong_ab604$wasted_DEM, na.rm = TRUE) /
  sum(votes_cong_ab604$DEM + votes_cong_ab604$REP, na.rm = TRUE)

```

```

tibble(Metric = c("Mean-Median Difference", "Efficiency Gap"),
       Value = c(mean_median_diff_ab604, efficiency_gap_ab604))

```

```

# A tibble: 2 × 2
  Metric      Value
  <chr>      <dbl>
1 Mean-Median Difference 0
2 Efficiency Gap      0.373

```