지방선거 예측 통계 보고서 (샘플)

Auto-generated

2025-09-12

목차

1 개요 1
1.1 재현 가능한 난수 시드 1

2 데이터 생성 1

3 그래프: 정당별 예측 득표율(난수 시뮬레이션) 3

4 gt 표: 동별 1·2위 및 격차 4

5 메모 6

1 개요

이 문서는 샘플 데이터를 **난수로 생성**하여 지방선거 득표 예측 절차를 시연합니다.
실제 예측에서는 공식 자료(NEC 등)로 대체하세요.

1.1 재현 가능한 난수 시드

2 데이터 생성

- 동별 유권자 수(2,000~6,000), 투표율(45%~75%)을 임의 생성
- 각 동에서 정당별 선호 비중을 임의로 생성 후 정규화

```
library(tidyverse)
library(gt)

#  /
cands <- c("  " = "  ", "  " = "  ", "  " = "  ")

n_precinct <- 25
base <- tibble(
```

```
    = paste0(" ", sprintf("%02d", 1:n_precinct)),
      = sample(2000:6000, n_precinct, replace = TRUE),
     = runif(n_precinct, 0.45, 0.75)
)

#      ( ) ->
shares <-
  base |>
  mutate(
     _raw = rexp(n_precinct, rate = 1.2),
     _raw = rexp(n_precinct, rate = 1.1),
     _raw = rexp(n_precinct, rate = 1.6)
  ) |>
  pivot_longer(ends_with("_raw"), names_to = " ", values_to = "raw") |>
  mutate(  = recode( ,
    " _raw" = "  ",
    " _raw" = "   ",
    " _raw" = "   "
  )) |>
  group_by( ) |>
  mutate(  = raw / sum(raw)) |>
  ungroup() |>
  select( ,  ,  )

#
votes <-
  shares |>
  left_join(base, by = " ") |>
  mutate(   = round(   *  ),
            = round(   *  )) |>
  select( ,  ,  ,  )

#    (    )
summary_party <-
  votes |>
  group_by( ) |>
  summarise(   = sum( ), .groups = "drop") |>
  mutate(   =    / sum( )) |>
  arrange(desc( ))

summary_party
```

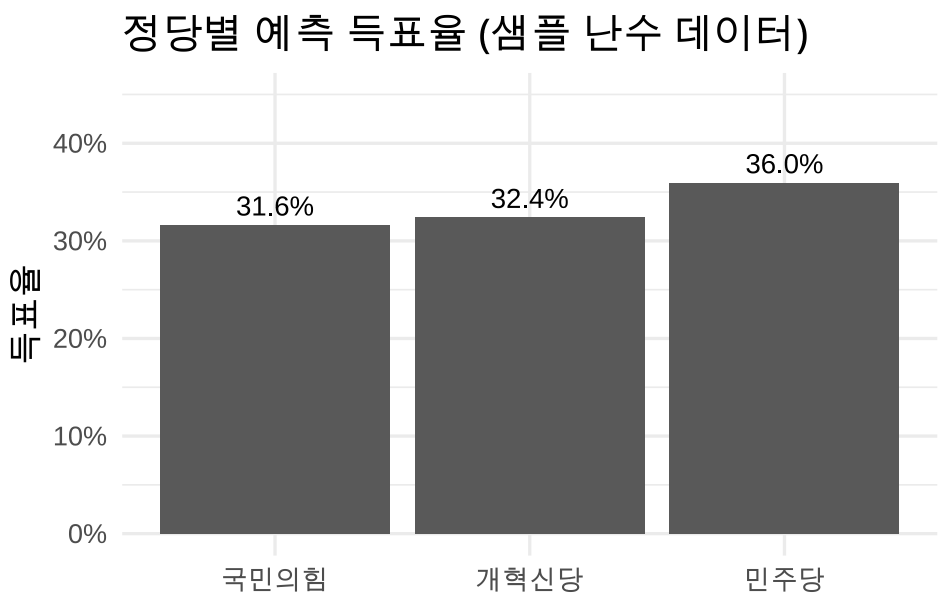
```
# A tibble: 3 x 3
```

```
  <chr>      <dbl>  <dbl>
1        23075  0.360
2      20793  0.324
3      20295  0.316
```

3 그래프: 정당별 예측 득표율(난수 시뮬레이션)

```
library(scales)
library(ggplot2)

ggplot(summary_party, aes(x = reorder( ,   ), y =   )) +
  geom_col() +
  geom_text(aes(label = percent(  , accuracy = 0.1)), vjust = -0.5, size = 4) +
  scale_y_continuous(labels = percent, limits = c(0, max(summary_party$  ) * 1.25))
  labs(title = "        (      )",
       x = NULL, y = "  ") +
  theme_minimal(base_size = 14)
```



해석(샘플): 난수 기반 시뮬레이션 결과로 특정 정당이 상대적으로 높은 득표율을 보일 수 있습니다. 이는 **데이터 생성 과정의 가정(선호 비중·투표율 분포)**에 좌우되며, 실제 분석에서는 **표본 설계·가중치·검증**을 통해 편향을 줄여야 합니다.

4 gt 표: 동별 1·2위 및 격차

아래 표는 동별 1위·2위 정당, 득표율 격차를 요약합니다.

```
#
by_precinct <- votes |>
  group_by( ) |>
  mutate(   =    / sum( )) |>
  arrange(desc( ), .by_group = TRUE) |>
  mutate(ranking = row_number()) |>
  ungroup()

top2 <- by_precinct |>
  dplyr::filter(ranking <= 2) |>
  select( ,  ,  ,  , ranking)

wide <- top2 |>
  pivot_wider(
    names_from = ranking,
    values_from = c( ,  ),
    names_sep = "_"
  ) |>
  mutate(  =  _1 -  _2) |>
  arrange(desc( ))

wide |>
  transmute(
    ,
    `1  ` =  _1,
    `1  ` =  _1,
    `2  ` =  _2,
    `2  ` =  _2,
    ` (%)` =  ,
    `  ` =
  ) |>
  gt() |>
  tab_header(
    title = md("**        (     )**"),
    subtitle = md("1·2     ,   ")
  ) |>
  fmt_percent(columns = c(`1  `, `2  `, ` (%)`), decimals = 1) |>
  fmt_number(columns = `  `, decimals = 0, use_seps = TRUE) |>
  cols_align(align = "center", columns = everything()) |>
  tab_options(table.font.size = px(14))
```

동별 상위 득표 요약 (샘플 난수 데이터)

1·2위 정당과 격차, 총투표수

동	1위 정당	1위 득표율	2위 정당	2위 득표율	격차(%)	총투표수
동07	개혁신당	88.9%	국민의힘	6.4%	82.5%	1,410
동11	국민의힘	86.5%	민주당	7.0%	79.4%	1,735
동21	민주당	86.6%	국민의힘	8.5%	78.2%	2,292
동24	민주당	85.1%	개혁신당	8.7%	76.4%	3,574
동04	개혁신당	82.4%	민주당	9.1%	73.3%	2,010
동12	개혁신당	79.3%	국민의힘	13.4%	65.9%	2,241
동08	민주당	79.4%	국민의힘	14.7%	64.7%	2,294
동15	개혁신당	74.4%	국민의힘	14.2%	60.1%	1,178
동23	국민의힘	71.2%	개혁신당	18.8%	52.4%	1,227
동10	국민의힘	72.6%	개혁신당	23.5%	49.1%	2,106
동01	국민의힘	72.5%	개혁신당	24.6%	48.0%	2,469
동03	민주당	71.8%	국민의힘	26.0%	45.9%	1,623
동14	개혁신당	60.6%	민주당	27.9%	32.6%	3,559
동06	개혁신당	60.2%	민주당	28.2%	31.9%	3,808
동18	민주당	52.8%	국민의힘	27.6%	25.2%	4,043
동25	민주당	58.6%	국민의힘	35.2%	23.4%	2,468
동09	개혁신당	49.2%	민주당	33.8%	15.5%	3,439
동02	국민의힘	47.3%	개혁신당	32.7%	14.6%	2,290
동20	국민의힘	55.2%	민주당	41.0%	14.2%	3,438
동05	민주당	41.2%	국민의힘	34.4%	6.8%	2,887
동13	국민의힘	40.2%	민주당	34.5%	5.7%	2,582
동22	국민의힘	37.8%	개혁신당	32.4%	5.4%	3,397
동17	국민의힘	44.1%	개혁신당	42.1%	2.0%	4,281
동19	민주당	48.4%	개혁신당	47.8%	0.6%	1,455
동16	민주당	44.5%	국민의힘	44.1%	0.3%	2,355

5 메모

- 본 문서는 **샘플 파이프라인**으로, 실제 예측에는 *공식 데이터 소스, 가중치 설계, 모델링(예: 다층 로지스틱/베이지안), 검증(홀드아웃/부트스트랩)*을 추가하세요.
- 그래프와 표는 `set.seed()`로 재현됩니다(시드 변경 시 결과가 달라짐).