

Male/Female priorities -REDS data

Indira Tirumala

2024-08-05

Goal: tabulate differences in ranking of issues between Gram Panchayats with women pradhan through reservation, versus GPs with open elections - for Rajasthan, UP and the rest of India. I.e. male/female priorities are not so different.

Data used: REDS, SEPRI 2014-2016.

Following Manu's email, I would like to mention (in our conclusion) the possibility that male/female priorities are not so different. So if you could tabulate the differences in Rajasthan, UP, and the rest of India, that would be helpful. We'll stick that table in the SI, but I can refer to it in the text. We'll want mean ratings for men and women, a column for the difference between the two, and a confidence interval around the difference. I can use the latter to say that, although the N is small, the range of the confidence interval is still narrow enough to rule out markedly different priorities.

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.3      v readr      2.1.4
## v forcats    1.0.0      v stringr   1.5.0
## v ggplot2    3.4.3      v tibble    3.2.1
## v lubridate  1.9.2      v tidyr     1.3.0
## v purrr      1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

## Warning: package 'kableExtra' was built under R version 4.3.2

##
## Attaching package: 'kableExtra'
##
## The following object is masked from 'package:dplyr':
##
##   group_rows

## # A tibble: 1 x 1
##   'n_distinct(village_name)'
##   <int>
## 1      189

## # A tibble: 1 x 1
##   'n_distinct(village_name)'
##   <int>
## 1      191
```

```

#Merging the datasets

final_data = inner_join(stateinfo, reservations, by = c("village_name", "gram_panchayat", "tehsil_taluka", "state"))

final_data = inner_join(final_data, rankings, by = c("village_name", "gram_panchayat", "tehsil_taluka", "state"))

## Warning in inner_join(final_data, rankings, by = c("village_name", "gram_panchayat", : Detected an unresolvable
## i Row 1 of 'x' matches multiple rows in 'y'.
## i Row 2117 of 'y' matches multiple rows in 'x'.
## i If a many-to-many relationship is expected, set 'relationship =
##   "many-to-many"' to silence this warning.

final_data %>% summarize(n_distinct(village_name))

## # A tibble: 1 x 1
##   'n_distinct(village_name)'
##               <int>
## 1                   189

final_data %>% count(state)

## # A tibble: 27 x 2
##   state          n
##   <chr>        <int>
## 1 ""           28
## 2 "0"           28
## 3 "ANDHRA PRADESH" 448
## 4 "BALODA BAZAR"   28
## 5 "BIHAR"         140
## 6 "CHHATISGARH"    84
## 7 "CHHATTISGARH"  112
## 8 "CHHTTISGARH"   28
## 9 "GUJARAT"       504
## 10 "GUJRAT"       56
## # i 17 more rows

final_data = final_data %>% mutate(state = recode(state,
  "TAAMIL NADU" = "TAMIL NADU",
  "TAMILNADU" = "TAMIL NADU",
  "TAMILNDU" = "TAMIL NADU",
  "BALODA BAZAR" = "CHHATTISGARH",
  "CHHATISGARH" = "CHHATTISGARH",
  "CHHTTISGARH" = "CHHATTISGARH",
  "HARIYANA" = "HARYANA",
  "MADHYA PRADESH" = "MADHYA PRADESH",
  "U.P" = "UTTAR PRADESH",
  "GUJRAT" = "GUJARAT",
  "MAHRASHTRA" = "MAHARASHTRA",
  "ODISHA 15" = "ODISHA"
))

final_data %>% count(state)

```

```
## # A tibble: 15 x 2
##   state      n
##   <chr>    <int>
## 1 ""        28
## 2 "0"       28
## 3 "ANDHRA PRADESH" 448
## 4 "BIHAR"    140
## 5 "CHHATTISGARH" 252
## 6 "GUJARAT"   560
## 7 "HARYANA"   252
## 8 "JHARKHAND"   70
## 9 "MADHYA PRADESH" 420
## 10 "MAHARASHTRA" 476
## 11 "ODISHA"    308
## 12 "RAJASTHAN"  642
## 13 "TAMIL NADU"  438
## 14 "UTTAR PRADESH" 812
## 15 "WEST BENGAL" 336
```

#Creating the table

```
table1 <- final_data %>%
  drop_na() %>%
  group_by(state, reserved_women, issues) %>%
  summarise(
    mean = mean(rank),
    sd = sd(rank),
    n = n()
  ) %>%
  pivot_wider(id_cols = c(state, issues),
              names_from = reserved_women, values_from = c(mean, sd, n)) %>%
  drop_na() %>%
  rename(reserved_women = mean_yes, open_election = mean_no) %>%
  mutate(difference = round(reserved_women - open_election, 1)) %>%
  mutate(se_diff = sqrt((sd_yes^2 / n_yes) + (sd_no^2 / n_no)),
         ci_lower = difference - qt(0.975, df = pmin(n_yes, n_no) - 1) * se_diff,
         ci_upper = difference + qt(0.975, df = pmin(n_yes, n_no) - 1) * se_diff) %>%
  mutate(ci_lower = round(ci_lower, 1),
         ci_upper = round(ci_upper, 1)) %>%
  mutate(CI_95 = paste(ci_lower, ci_upper, sep = " , ")) %>%
  select(!sd_yes:sd_no) %>%
  select(!se_diff:ci_upper) %>%
  mutate(reserved_women = round(reserved_women, 1),
         open_election = round(open_election, 1)) %>%
  rename(n_reserved = n_yes) %>%
  rename(n_open = n_no) %>%
  ungroup()
```

'summarise()' has grouped output by 'state', 'reserved_women'. You can override
using the 'groups' argument.

#Creating table for Rajasthan only

```

rajs_table = table1 %>%
  filter(., state=="RAJASTHAN") %>%
  mutate(Number_of_Villages = n_reserved + n_open) %>%
  rename(Issues = issues) %>%
  rename(Reserved_for_women = reserved_women) %>%
  rename(Open_election = open_election) %>%
  rename(Difference = difference) %>%
  rename(Confidence_Interval_95 = CI_95) %>%
  select(!c(state, n_reserved, n_open))

```

```

#Outputting the table for Rajasthan

```

```

#kable(rajs_table, format = "latex", booktabs = TRUE, caption = "Caption to be inserted here") %>%
  #cat()

```