

MEDS Hiring Exercise

Salma Khalfallah

2026-02-12

```
# prep work
knitr::opts_chunk$set(echo = TRUE)
# setting seed for reproducibility
set.seed(1)
# importing necessary packages:
library(tidyverse)

## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## vforcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.2     v tibble    3.3.0
## v lubridate 1.9.4     v tidyr    1.3.1
## v purrr    1.1.0
## -- 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

library(stringr)
```

Data Wrangling

1.

```
## PART ONE: Wrangling

part1_data <- read.csv("part1_data.csv")
part1_fips <- read.csv("part1_fips.csv")

# merging with fips dataset
part1_data_merged <- merge(part1_data, part1_fips, by.x = "ReportingCountyName", by.y = "county_name")

# identifying smallest election reporting unit aka precincts
# notice: the smallest electoral units contain numbers next to the township name
# we can use this

popular_parties <- c("DEMOCRAT", "REPUBLICAN", "INDEPENDENT", "LIBERTARIAN")

part1_cleaned <- part1_data_merged |>
  select("DataEntryJurisdictionName", "Office", "PoliticalParty", "TotalVotes", "ReportingCountyName",
```

```

    "NameonBallot", "Election") |>
filter(str_detect(DataEntryJurisdictionName, "\\d") == TRUE) |>
mutate(Office = recode(Office,
                        "US President & Vice President" = "US PRESIDENT")) |>
rename("office" = Office) |>
rename("party_detailed" = PoliticalParty) |>
rename("precinct" = DataEntryJurisdictionName) |>
mutate(party_detailed = recode(party_detailed,
                                "Democratic" = "DEMOCRAT",
                                "Republican" = "REPUBLICAN",
                                "Independent" = "INDEPENDENT",
                                "Libertarian" = "LIBERTARIAN",
                                .default = str_to_upper(party_detailed)
)) |>
mutate(party_simplified = if_else(party_detailed %in% popular_parties,
                                   party_detailed,
                                   "OTHER")) |>
rename("votes" = TotalVotes) |>
mutate(votes = as.numeric(votes)) |>
rename("county_name" = ReportingCountyName) |>
mutate(jurisdiction_name = county_name) |>
mutate(write_in = str_detect(NameonBallot, "W/I")) |>
mutate(candidate = str_split(NameonBallot, " & ", simplify = TRUE)[,1]) |>
mutate(candidate = str_remove(candidate, "\\(W/I\\)")) |>
mutate(candidate = str_to_upper(candidate)) |>
mutate(year = "2024") |>
mutate(stage = "GEN") |>
mutate(special = FALSE) |>
mutate(date = as.Date("2024-11-05")) |>
select(-NameonBallot, -Election)

```

2. There are some issues with the data with respect to data standardization and coding practices. For example, instead of there being a boolean function for the write-in function, the “write-in” indication is within the string for the named candidate and is not immediately obvious what the write-in indicator for the set is. During future data entry, standard data practices should be implemented to prevent these kind of issues.

3.

```

county_totals <- part1_cleaned |>
group_by(county_name, candidate) |>
summarise(total_votes = sum(votes, na.rm = TRUE)) |>
ungroup()

```

```

## `summarise()` has grouped output by 'county_name'. You can override using the
## `groups` argument.

```

```

knitr::kable(county_totals)

```

county_name	candidate	total_votes
Brookside County	ANDRÉ RAMON MCNEIL SR.	0
Brookside County	CAROL ASHER	0
Brookside County	CHASE OLIVER	157
Brookside County	CHERUNDA LYNN FOX	0
Brookside County	CLAUDIA DE LA CRUZ	3
Brookside County	CORNEL WEST	0
Brookside County	DONALD J TRUMP	20253
Brookside County	DOUG JENKINS	0
Brookside County	KAMALA D. HARRIS	5183
Brookside County	NALA BAOZUN SCOTT JOHNSON JR	0
Brookside County	PAIJ BORING	0
Brookside County	PETER SONSKI	5
Brookside County	ROBERT F. KENNEDY JR.	239
Brookside County	SHONDRA IRVING	0
Brookside County	SUSAN MAUDE BUCHSER	0
Brookside County	TOM HOEFLING	2
Brookside County	WILLIAM CODY NALBACH	0
Cedar County	ANDRÉ RAMON MCNEIL SR.	0
Cedar County	CAROL ASHER	0
Cedar County	CHASE OLIVER	283
Cedar County	CHERUNDA LYNN FOX	0
Cedar County	CLAUDIA DE LA CRUZ	6
Cedar County	CORNEL WEST	1
Cedar County	DONALD J TRUMP	20699
Cedar County	DOUG JENKINS	0
Cedar County	JØHN\$\$ SMÎT#H	0
Cedar County	KAMALA D. HARRIS	11947
Cedar County	NALA BAOZUN SCOTT JOHNSON JR	0
Cedar County	PAIJ BORING	0
Cedar County	PETER SONSKI	13
Cedar County	ROBERT F. KENNEDY JR.	393
Cedar County	SHONDRA IRVING	0
Cedar County	SUSAN MAUDE BUCHSER	0
Cedar County	TOM HOEFLING	2
Cedar County	WILLIAM CODY NALBACH	0
Clearwater County	ANDRÉ RAMON MCNEIL SR.	0
Clearwater County	CAROL ASHER	0
Clearwater County	CHASE OLIVER	45
Clearwater County	CHERUNDA LYNN FOX	0
Clearwater County	CLAUDIA DE LA CRUZ	1
Clearwater County	CORNEL WEST	0
Clearwater County	DONALD J TRUMP	6986
Clearwater County	DOUG JENKINS	0
Clearwater County	KAMALA D. HARRIS	1443
Clearwater County	NALA BAOZUN SCOTT JOHNSON JR	0
Clearwater County	PAIJ BORING	0
Clearwater County	PETER SONSKI	1
Clearwater County	ROBERT F. KENNEDY JR.	64
Clearwater County	SHONDRA IRVING	0
Clearwater County	SUSAN MAUDE BUCHSER	0
Clearwater County	TOM HOEFLING	0
Clearwater County	WILLIAM CODY NALBACH	0

county_name	candidate	total_votes
Fairview County	ANDRÉ RAMON MCNEIL SR.	0
Fairview County	CAROL ASHER	0
Fairview County	CHASE OLIVER	42
Fairview County	CHERUNDA LYNN FOX	0
Fairview County	CLAUDIA DE LA CRUZ	0
Fairview County	CORNEL WEST	0
Fairview County	DONALD J TRUMP	7779
Fairview County	DOUG JENKINS	0
Fairview County	KAMALA D. HARRIS	2084
Fairview County	NALA BAOZUN SCOTT JOHNSON JR	0
Fairview County	PAIJ BORING	0
Fairview County	PETER SONSKI	0
Fairview County	ROBERT F. KENNEDY JR.	87
Fairview County	SHONDRA IRVING	0
Fairview County	SUSAN MAUDE BUCHSER	0
Fairview County	TOM HOEFLING	0
Fairview County	WILLIAM CODY NALBACH	0
Heritage County	ANDRÉ RAMON MCNEIL SR.	0
Heritage County	CAROL ASHER	0
Heritage County	CHASE OLIVER	311
Heritage County	CHERUNDA LYNN FOX	0
Heritage County	CLAUDIA DE LA CRUZ	23
Heritage County	CORNEL WEST	7
Heritage County	DONALD J TRUMP	26067
Heritage County	DOUG JENKINS	0
Heritage County	KAMALA D. HARRIS	19017
Heritage County	NALA BAOZUN SCOTT JOHNSON JR	0
Heritage County	PAIJ BORING	0
Heritage County	PETER SONSKI	18
Heritage County	ROBERT F. KENNEDY JR.	510
Heritage County	SHONDRA IRVING	0
Heritage County	SUSAN MAUDE BUCHSER	0
Heritage County	TOM HOEFLING	0
Heritage County	WILLIAM CODY NALBACH	0
Mountain County	ANDRÉ RAMON MCNEIL SR.	0
Mountain County	CAROL ASHER	0
Mountain County	CHASE OLIVER	32
Mountain County	CHERUNDA LYNN FOX	0
Mountain County	CLAUDIA DE LA CRUZ	0
Mountain County	CORNEL WEST	0
Mountain County	DONALD J TRUMP	3609
Mountain County	DOUG JENKINS	0
Mountain County	KAMALA D. HARRIS	1219
Mountain County	NALA BAOZUN SCOTT JOHNSON JR	0
Mountain County	PAIJ BORING	0
Mountain County	PETER SONSKI	0
Mountain County	ROBERT F. KENNEDY JR.	74
Mountain County	SHONDRA IRVING	0
Mountain County	SUSAN MAUDE BUCHSER	0
Mountain County	TOM HOEFLING	0
Mountain County	WILLIAM CODY NALBACH	0
Northridge County	ANDRÉ RAMON MCNEIL SR.	1

county_name	candidate	total_votes
Northridge County	CAROL ASHER	1
Northridge County	CHASE OLIVER	409
Northridge County	CHERUNDA LYNN FOX	1
Northridge County	CLAUDIA DE LA CRUZ	21
Northridge County	CORNEL WEST	4
Northridge County	DONALD J TRUMP	45652
Northridge County	DOUG JENKINS	1
Northridge County	KAMALA D. HARRIS	23082
Northridge County	NALA BAOZUN SCOTT JOHNSON JR	1
Northridge County	PAIJ BORING	1
Northridge County	PETER SONSKI	34
Northridge County	ROBERT F. KENNEDY JR.	551
Northridge County	SHONDRA IRVING	1
Northridge County	SUSAN MAUDE BUCHSER	1
Northridge County	TOM HOEFLING	3
Northridge County	WILLIAM CODY NALBACH	1
Oakmont County	ANDRÉ RAMON MCNEIL SR.	0
Oakmont County	CAROL ASHER	0
Oakmont County	CHASE OLIVER	85
Oakmont County	CHERUNDA LYNN FOX	0
Oakmont County	CLAUDIA DE LA CRUZ	1
Oakmont County	CORNEL WEST	0
Oakmont County	DONALD J TRUMP	9108
Oakmont County	DOUG JENKINS	0
Oakmont County	KAMALA D. HARRIS	3135
Oakmont County	NALA BAOZUN SCOTT JOHNSON JR	0
Oakmont County	PAIJ BORING	0
Oakmont County	PETER SONSKI	6
Oakmont County	ROBERT F. KENNEDY JR.	115
Oakmont County	SHONDRA IRVING	0
Oakmont County	SUSAN MAUDE BUCHSER	0
Oakmont County	TOM HOEFLING	0
Oakmont County	WILLIAM CODY NALBACH	0
Prairie County	ANDRÉ RAMON MCNEIL SR.	0
Prairie County	CAROL ASHER	0
Prairie County	CHASE OLIVER	45
Prairie County	CHERUNDA LYNN FOX	0
Prairie County	CLAUDIA DE LA CRUZ	0
Prairie County	CORNEL WEST	0
Prairie County	DONALD J TRUMP	2922
Prairie County	DOUG JENKINS	0
Prairie County	KAMALA D. HARRIS	1459
Prairie County	NALA BAOZUN SCOTT JOHNSON JR	0
Prairie County	PAIJ BORING	0
Prairie County	PETER SONSKI	3
Prairie County	ROBERT F. KENNEDY JR.	35
Prairie County	SHONDRA IRVING	0
Prairie County	SUSAN MAUDE BUCHSER	0
Prairie County	TOM HOEFLING	0
Prairie County	WILLIAM CODY NALBACH	0
Willow County	ANDRÉ RAMON MCNEIL SR.	0
Willow County	CAROL ASHER	0

county_name	candidate	total_votes
Willow County	CHASE OLIVER	40
Willow County	CHERUNDA LYNN FOX	0
Willow County	CLAUDIA DE LA CRUZ	0
Willow County	CORNEL WEST	0
Willow County	DONALD J TRUMP	3481
Willow County	DOUG JENKINS	0
Willow County	KAMALA D. HARRIS	1204
Willow County	NALA BAOZUN SCOTT JOHNSON JR	0
Willow County	PAIJ BORING	0
Willow County	PETER SONSKI	0
Willow County	ROBERT F. KENNEDY JR.	54
Willow County	SHONDRA IRVING	0
Willow County	SUSAN MAUDE BUCHSER	0
Willow County	TOM HOEFLING	0
Willow County	WILLIAM CODY NALBACH	0