Canadian Federal Election

Geunchul Shin

#Part 1

```
#### Preamble ####
# Purpose: Get data on 2021 Canadian Federal election
# Author: Geunchul Shin
# Email: joey.shin@mail.utoronto.ca
# Github: (https://github.com/geunchulshin/Tutorial-1)
# Date: 9 January 2024
```

```
#### Workspace setup ####
#install.packages("tidyverse")
#install.packages("janitor")

library(tidyverse)
```

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.4
v forcats   1.0.0     v stringr   1.5.1
v ggplot2   3.4.4     v tibble    3.2.1
v lubridate 1.9.3     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 beco
```

```
library(janitor)
```

```
Attaching package: 'janitor'

The following objects are masked from 'package:stats':

    chisq.test, fisher.test
```

```
simulated_elections_data <-
  tibble(
    riding = rep(c(1:338)),
    elected_party= sample(c('Liberal','Conservative','Bloc Québécois','New Democratic','Gr
  )
```

#Part 2

```
raw_elections_data <-
  read_csv(
    file =
      "https://www.elections.ca/res/rep/off/ovr2021app/53/data_donnees/table_tableau11.csv
    show_col_types = FALSE
  )

write_csv(x = raw_elections_data, file = "election.csv")
```

```
cleaned_elections_data <-
  clean_names(raw_elections_data)
```

```
cleaned_elections_data <-
  cleaned_elections_data |>
  select(electoral_district_name_nom_de_circonscription, elected_candidate_candidat_elu)
```

```
cleaned_elections_data <-
  cleaned_elections_data |>
  rename(
    electoral_district_name = electoral_district_name_nom_de_circonscription,
    elected_candidate = elected_candidate_candidat_elu)
```

```
cleaned_elections_data <-
  cleaned_elections_data |>
```

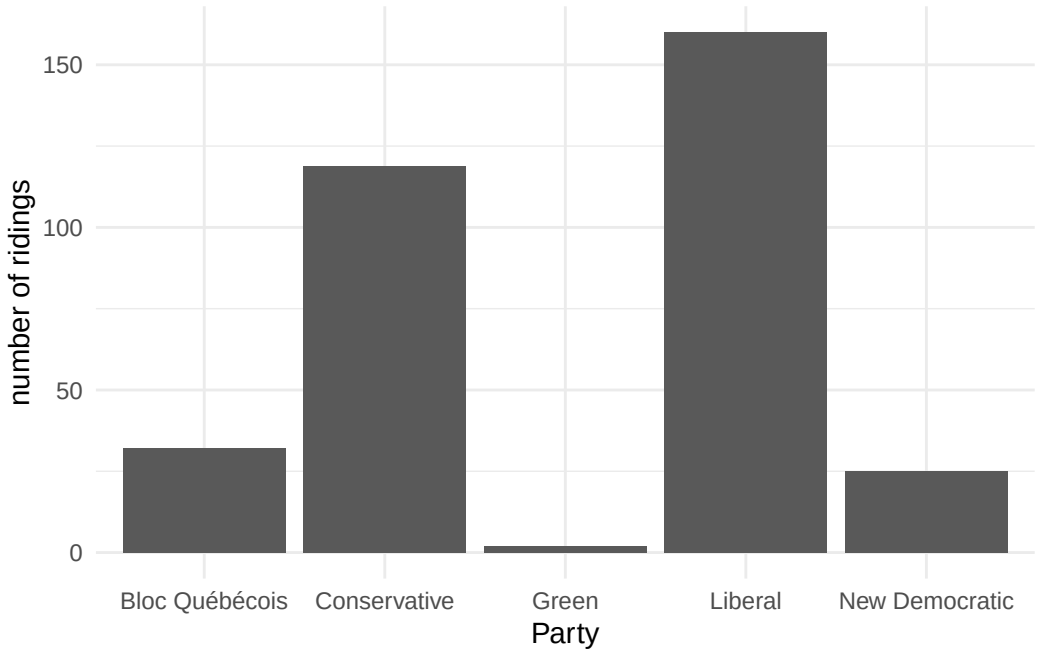
```
  separate(
    col = elected_candidate,
    into = c("Other", "party"),
    sep = "/"
  ) |>
  select(-Other)
```

```
cleaned_elections_data$party |>
  unique()
```

```
[1] "Libéral"                          "Conservateur"
[3] "Bloc Québécois"                   "NPD-Nouveau Parti démocratique"
[5] "Parti Vert"
```

```
cleaned_elections_data <-
  cleaned_elections_data |>
  mutate(
    party =
      case_match(
        party,
        "Libéral" ~ "Liberal",
        "Conservateur" ~ "Conservative",
        "Bloc Québécois" ~ "Bloc Québécois",
        "NPD-Nouveau Parti démocratique" ~ "New Democratic",
        "Parti Vert" ~ "Green"
      )
  )
```

```
cleaned_elections_data |>
  ggplot(aes(x=party)) +
  geom_bar() +
  theme_minimal() +
  labs(x="Party", y="number of ridings")
```


150
100
50
0
number of ridings
Bloc Québécois
Conservative
Green
Liberal
New Democratic
Party