````
---
title: "Exploring the Canadian Election"
author: "Luca Carnegie"
format: html
editor: visual
---

```{r}
#### Preamble #
# Purpose: Read in data from Elections Canada and see how many seats each party won
# Author: Luca Carnegie
# Email: luca.carnegie@mail.utoronto.ca
# Date: 10 Jan 2023
# Prerequisites: The Australian example
```

```{r}
#### Workspace setup ####
# install.packages("tidyverse")
# install.packages("janitor")
# install.packages("gtranslate")

library(tidyverse)
library(janitor)
library(gtranslate)

```

```{r}
# Simulated Data
simulated_data <-
  tibble(
    # Use 1 through to 151 to represent each division
    "Division" = 1:338,
    # Randomly pick an option, with replacement, 338 times
    "Party" = sample(
      x = c("Liberal", "Conservative", "New Democratic", "Green", "Bloc Québécois",
"Other"),
      size = 338,
      replace = TRUE
    )
  )

simulated_data
```

```{r}
#Real Data
## Read in Elections Canada Data ##
raw_elections_data <-
  read_csv(
    file =
"https://www.elections.ca/res/rep/off/ovr2021app/53/data_donnees/table_tableau11.csv",
    show_col_types = FALSE,
  )

#Save elections data locally
write_csv(
  x = raw_elections_data,
  file = "cad_voting.csv"
)
```
```{r}
## Clean Data ##
````
````

```
# Basic cleaning #
raw_elections_data <-
  read_csv(
    file = "cad_voting.csv",
    show_col_types = FALSE
  )

# Make the names easier to type
cleaned_elections_data <-
  clean_names(raw_elections_data)

#Make the names shorter
cleaned_elections_data <-
  cleaned_elections_data |>
  rename(
    electoral_district = electoral_district_name_nom_de_circonscription,
    elected_candidate = elected_candidate_candidat_elu
  )

#Get the winning party, in French
cleaned_elections_data <-
  cleaned_elections_data |>
  separate(
    col = elected_candidate,
    into = c("Other", "party"),
    sep = "/"
  ) |>
  select(-Other)

#cleaned_elections_data$party |>
 # unique()

#Translate the column back to English
cleaned_elections_data <-
  cleaned_elections_data |>
  mutate(
    elected_party =
      case_match(
        party,
        "Libéral" ~ "Liberal",
        "Conservateur" ~ "Conservative",
        "NPD-Nouveau Parti démocratique" ~ "NDP",
        "Parti Vert" ~ "Green",
        "Bloc Québécois" ~ "Bloc Québécois"
      )
  )

cleaned_elections_data

#save the cleaned data

write_csv(
  x = cleaned_elections_data,
  file = "cleaned_elections_data.csv"
)
```


```{r}
## Exploring the Data ##

# Read the data #
cleaned_elections_data <-
  read_csv(
```

```
    file = "cleaned_elections_data.csv",
    show_col_types = FALSE
  )

#Seats each party Won
cleaned_elections_data |>
  count(elected_party)

#Make graph with ggplot
cleaned_elections_data |>
  ggplot(aes(x = elected_party)) +
  geom_bar() +
  theme_minimal() + # Make the theme neater
  labs(x = "Party", y = "Number of seats") # Make labels more meaningful
```