Facebook Ads in 2019 Canada Federal Election Reveal Limitation of Female Participation in Politics in Canada

Abstract

Is politic still a good old boy's club in Canada? For some parties in Canada, it might still is. In this paper, the audience demography of 14690 Facebook ads plublised during the election season were analysed. The findings indicates that three out of the six parties who participated in the Federal Election have their ads reached substaintially more male than female audience. Moreover, it is also found that most of the parties have different messaging for female and male audiences — environmental and social topics for females, and economic topics for males.

1. Introduction

The 2019 Canadian Federal Election was held on October 21, 2019, to elect members of the House of Commons to the 43rd Canadian Parliament. Six parties—namely Bloc Québécois, Conservative Party, Green Party, Liberal Party, New Democratic Party (NDP), and People's Party of Canada (PPC)—were competing for the seats. As a result, the Liberal Party of Canada was returned with a minority of the seats, and the Conservative Party came second. Bloc Quebecois came third with NDP came forth. The Green party is the 5th and PPC had won no seat (The Guardian 2019). During the election season, Facebook ads had been adapted by all candidate parties as a campaigning platform. We collected the Facebook ads that were published by the six parties during the election season and analyze the distribution of the ads in terms of region, gender and age groups.

In March 2019, Facebook launched the Ad Library — a direct response to the Cambridge Analytica scandal that hit the platform in early 2018 (Silva et al. 2020). The Facebook Ad Library is a place where anyone can see ads placed by any Facebook page, and the information can be retrieved through the provided API. Through the API, we are able to see who pays for an ad, the gender and age group distributions reached by the ad, the dates the ad is published and other information. Through the analysis in gender distribution of all parties ads, we found that although both male and female (and all genders) citizens have the right to vote in Canada, the equal inclusion of the female audience in election campaigns is not reflected in all parties' targeting strategies. The findings indicate that three out of the six parties (Bloc Québécois, Conservative Party, and PPC) have their ads reached more female audience than male audience, despite that there were more female users than male users on Facebook in Canada.

In this work, we examine the dataset of advertisements posted to Facebook at the time of the 2019 Canada federal election season by exploratory data analysis, sentiment analysis and TF-IDF keyword extraction. Besides the gender distribution of the ads, we also further answer the following questions:

- How much money have the parties spent on Facebook advertisements during the election season?
- What is the regional distribution and age group distribution of the ads?
- What is the overall sentiment of the ads?
- Finally, what are the narratives of the ads and do they differ by gender?

2. Data

The data is collected through Facebook Ad Library Scraper developed by (Woolf (n.d.)). The queries are set to collect the ads from the pages of each party and each leader that were active from June 1st, 2019 to October 21, 2019. Three data frames are retrieved — a data frame that contains the raw ads and their metadata; a data frame that has the unnested demographic distributions of people reached by ads, and the data frame that has the unnested region distributions of people reached by ads. The three data frames can be joined by the unique identifier of the ads.

For each of the three data frames, the ads from each party page and the ads from each party leader page are grouped as ads from one party. For example, ads from page "Liberal Party of Canada | Parti libéral du Canada" and page "Justin Trudeau" are grouped as ads from Liberal Party.

3. Results

3.1 Overall Spending

Table 1 displays the overall spending and impressions of each party. As the two biggest political parties in Canada, both Liberal Party and Conservative Party had spent more than 1 million for Facebook ads during the election season between June 1st, 2019 to October 21, 2019. The Liberal Party has put 12199 ads, which is the largest amount of ads among all parties, and six times more than the Conservative Party. It is estimated that Liberal Party has spent at least more than 1.7 million on Facebook ads during the election season, and the ads had made at least more than 185 million impressions. The regional and demographic distribution of people reached by the ads is examined in the following sections.

Table 1: Ad counts, spending and impressions of each party

Party	Ad Counts	Min Spending	Max Spending	Min Impressions	Max Impressions
Liberal	12199	1766600	3254501	162004000	185655813
Conservative	2025	1370300	1841075	105810000	102794994
NDP	312	417900	532888	27839000	28004692
Bloc Québécois	44	26700	35056	3799000	4495956
Green	44	50100	62356	3042000	3653956
PPC	66	19400	28334	2352000	2799934

3.2 Audience Targeting

Spending by Regions

The regional distribution of ad spending in each party is shown in Figure 1. Most of the parties had their ads reached the most people in Ontario except for Bloc Québécois who reached the most people in Quebec (99.99%) and Green Party who reached the most people in British Columbia (67.73%). The ads of the Liberal Party, Conservative Party, NDP and Green Party even reached a small percentage of people in the United States. It is worth noting that although 14.79% of the spending of PPC is in Quebec, there is no french creative body found in PPC's ads.

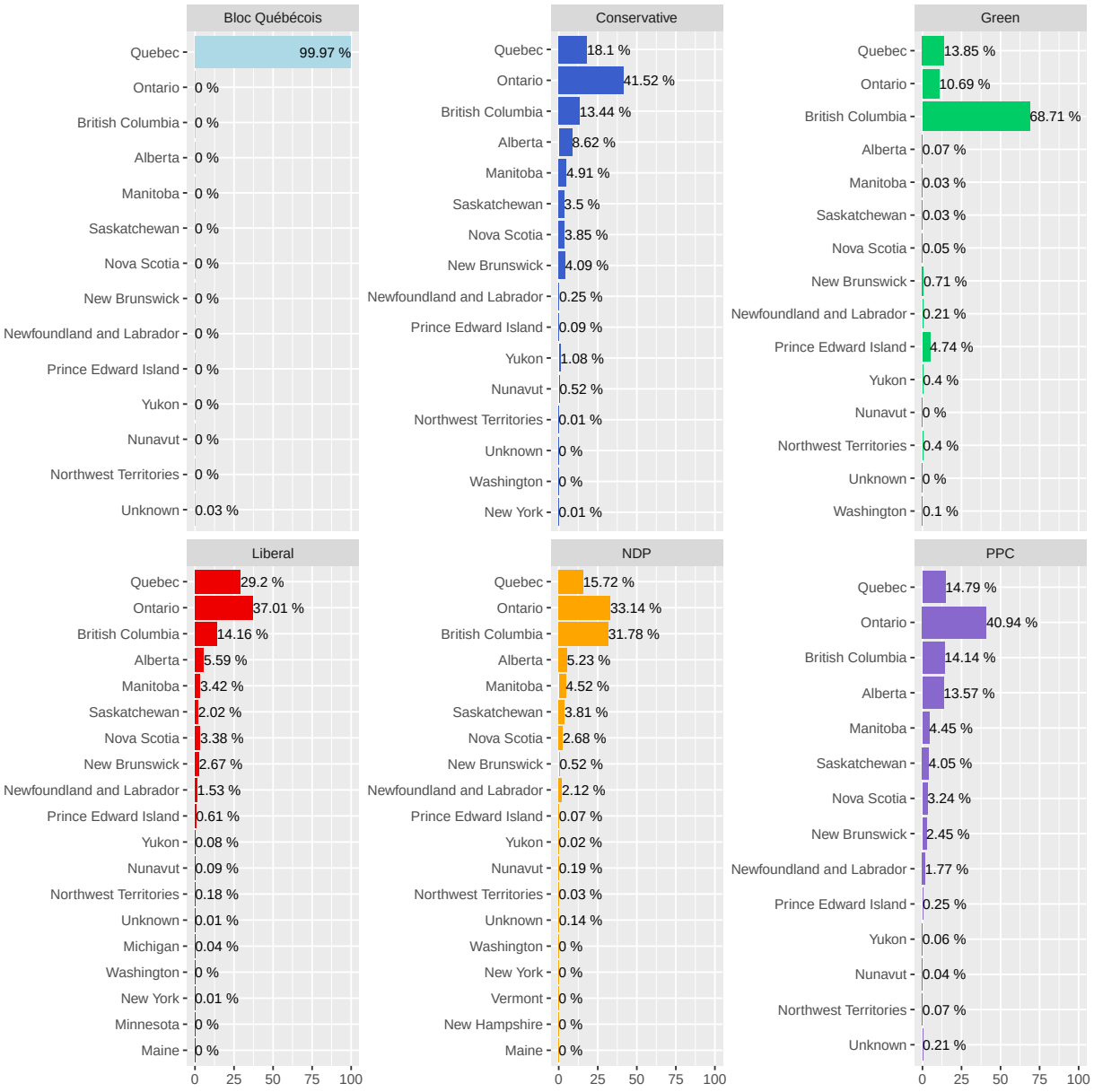


Figure 1: Reginal distribution of ads spending in each party

Spending by Gender

According to (("Statista" 2021)), at the time of the election season (October 2019), women accounted for 53.2 percent of Facebook users in Canada. The gender distribution of the ads does not reflect the gender ratio of Facebook users in Canada, as shown in Figure 2. PPC's ads spending demonstrated the largest proportion of male audience; 74.73% of its ad spending has been distributed to reach male audiences. The ads of Conservative Party and Bloc Québécois also had reached more males than females. The Green Party, on the other hand, has 64.45% of ad spending for reaching female audiences, which is the highest among all parties.

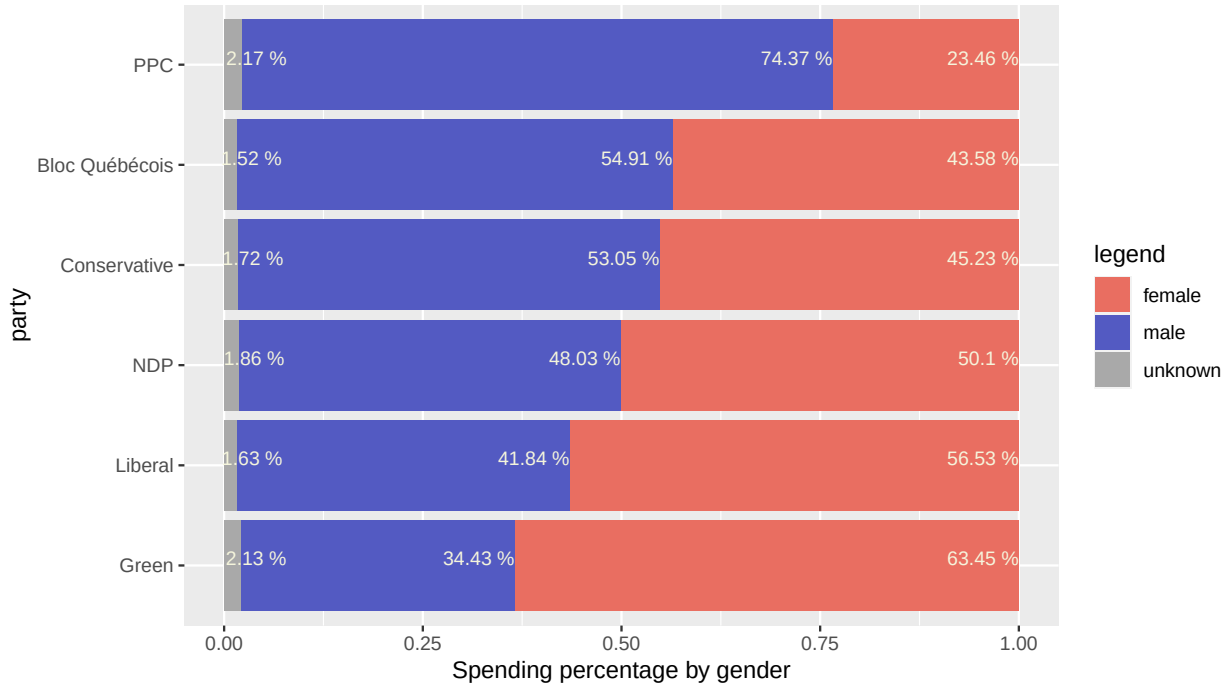


Figure 2: Gender distribution of ads spending in each party

Each ad has its own gender distribution. If an ad has reached more female audiences than male audiences, it is labelled as a "more-female-reaching" ad and vice versa. The number of "more-female-reaching" ads and "more-male-reaching" are counted for each party, which is presented in Table 2. The proportions are then calculated to have an idea if a party has extremely skewed gender distribution in their ads; the result is presented in Table 3. The results show that PPC, Bloc Québécois and Conservative Party have a substantially large portion of ads that reached more male audiences than female audiences. Out of PPC's 66 ads, only 1 ad reached more female audience than male audience. For Bloc Bloc Québécois, more-female-reaching ads only accounted for 10% of all their ads; only one-third of Conservative Party's ads reached more female audience than male audience.

Table 2: Number of more-female-reaching ads and more-male-reaching ads

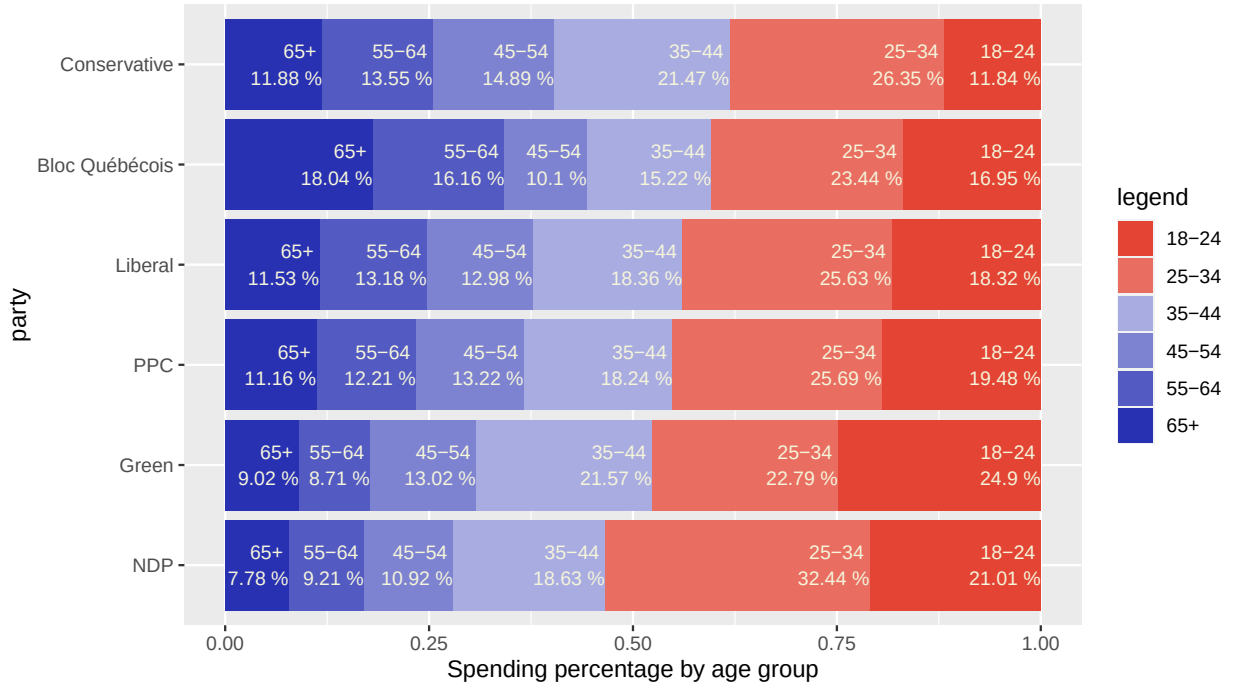
	Bloc Québécois	Conservative	Green	Liberal	NDP	PPC
Ads reached more females	4	659	40	8215	167	1
Ads reached more males	40	1366	4	3984	145	65

Table 3: Proprotions of more-female-reaching ads and more-male-reaching ads

	Bloc Québécois	Conservative	Green	Liberal	NDP	PPC
Ads reached more females	0.09	0.33	0.91	0.67	0.54	0.02
Ads reached more males	0.91	0.67	0.09	0.33	0.46	0.98

Spending by Age Groups

According to (("Statista" 2021)), in October 2019, 43.9% of the Facebook users in Canada were below 35-year-old and 56.1% were 35-year-old or older. Besides NDP, the ads of all parties had reached more people who are 35-year-old or older. NDP's ads had 53.45% of their ads spending contributed to reaching younger generations. Results are shown in Figure 3.



3.3 Sentiments of the ads

Although it is a common tactic to attack rivals and deliver negative comments of the status quo, the ads in each party remain more positive than negative. PPC's ads have the highest average sentiment score. Among all parties, Conservative Party has the widest range of sentiment scores for their ads. The lowest sentiment score of all the ads is -0.85, and it belongs to the Conservative Party shown in Figure 5. The ad attacks its rival by the association of "gun" and "gangster", which evokes emotions of fear.

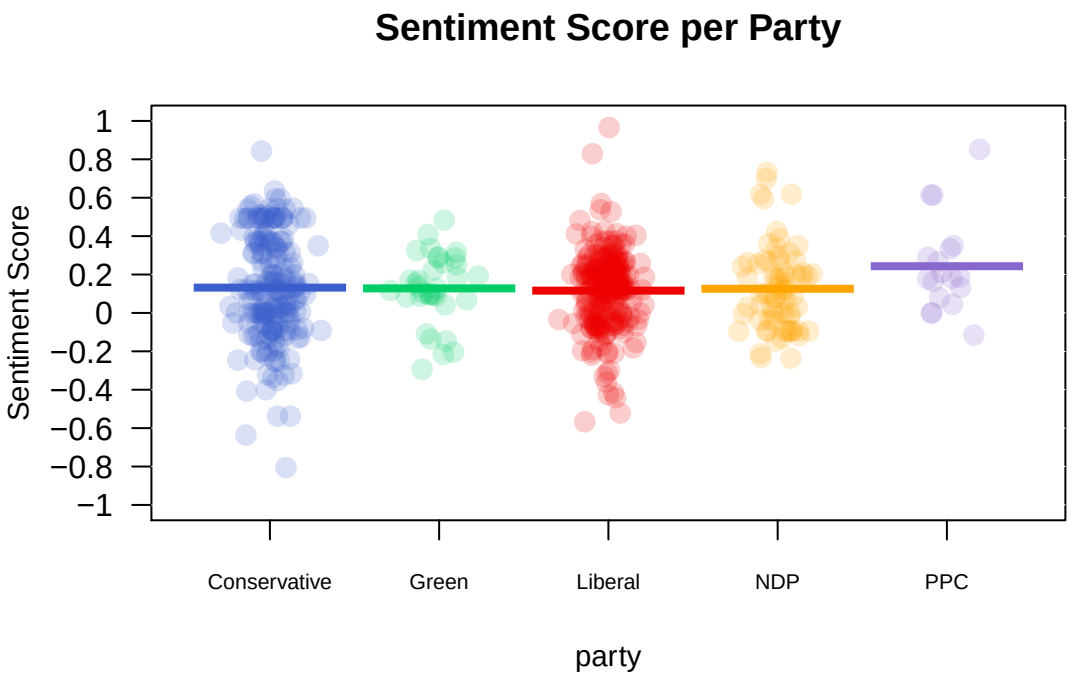


Figure 3: Average sentiment score of ads in each party

3.4 TF-IDF Keywords

The results of TF-IDF of each party's ads, displayed in Figure 6, show the focused narratives of the parties by the extracted keywords. These keywords also indirectly reveal the characteristics of each party. Bringing back "public" "transit" "tax" "credit" was the main manifesto of the Conservative Party; Green Party has their focus on environmental issues such as "fracking"; Liberal Party has focused on "grassroots" "fundraising" and also "middle" "class" related issues; NDP has their focus on "education", "health care", "housing", and also wealth distribution; PPC stressed Canadian values and identity along with the mentions of "veterans", ending official "multiculturalism", and "freedom". Data of Bloc Québécois is missing because their ads are in French only.

To further investigate in the differences messaging by gender, TF-IDF is conducted to extract keywords from more-female-reaching ads and more-male-reaching-ads respectively. The results are shown in Figure 7. For some parties, the segmentation of gender is obvious. For example, although there are overlapped keywords such as "tax" and "ahead", Conservative Party stressed "maternity" benefits and helping "parents" in their more-female-reaching ads. In Green Party's more-male-reaching ads, "Elizabeth May", which is Green Party's female leader was not on the keyword list; and instead of environmental issues such as "fracking", they focus more on the economical keywords such as "economy" and "workers". Liberal Party's

Figure 4: The most negative ad has sentiment score -.85 and it belongs to the Conservative Party

more-female-reaching ads stressed on strengthened gun controls ("ban" "military-style" "assault" "rifles" "restrict" "handguns"); their more-male-reaching ads seemed to focus on economic topic as well ("middle" "class"). NDP's ads also have different messaging towards different genders. Their more-female-reaching ads have distinctive keywords "education", "healthcare" and "housing", whereas their more-male-reaching ads have "corporations", "telecom" and "governments". The keywords of PPC's more-female-reaching ad were generated from only one ad; the original creative body for this ad is "Watch our Veterans Policy announcement in Gatineau with 42 PPC veterans on stage. #PPC2019". Since there is only one more-female-reaching ad, it is not clear if PPC had crafted the ads to deliver different messaging for female audiences.

4. Discussion and Conclusion

Is Canada's politics still an old boy club? For some parties in Canada, it might still is. From the observations in gender and age group distribution, we found that three out of the six parties' ads have reached older generations more and male audiences more. These parties, namely Conservative, PPC and Bloc Québécois, had targeting strategies that resulted in reaching substantially more male audiences than female audiences. Because how interests or behaviours of the audience are targeted is not disclosed, we can not know exactly what had led to the larger portion of male audiences reached by the ads of certain parties. However, we argue that their targeting strategy can indirectly reflect their hidden mentality towards women's participation in politics.

Through TF-IDF keyword extraction, we also found that the ads have the tendency to speak about economy-related issues to men and environmental and societal issues to women. In the ads of the Liberal Party, Conservative Party, Green Party and NDP, the more-female-reaching ads are surrounding issues of fracking, gun control, education, and healthcare; for their more-male-reaching-ads, the ads are centring around the economy, workers, middle-class growth, and corporations. On the other hand, Conservative Party had the focus on family and maternity benefits for their female audience. Because PPC only had one more-female-reaching ad, there is no discernible topic for their female audience. This gender bias in keywords and narratives might occur naturally also due to targeting specific interests that influenced the gender distribution of the people reached. However, it can be argued that women are expected to "care" about politics for only certain concerns.

In conclusion, we argue that the gender bias in distribution and messaging of the 2019 Canadian Federal

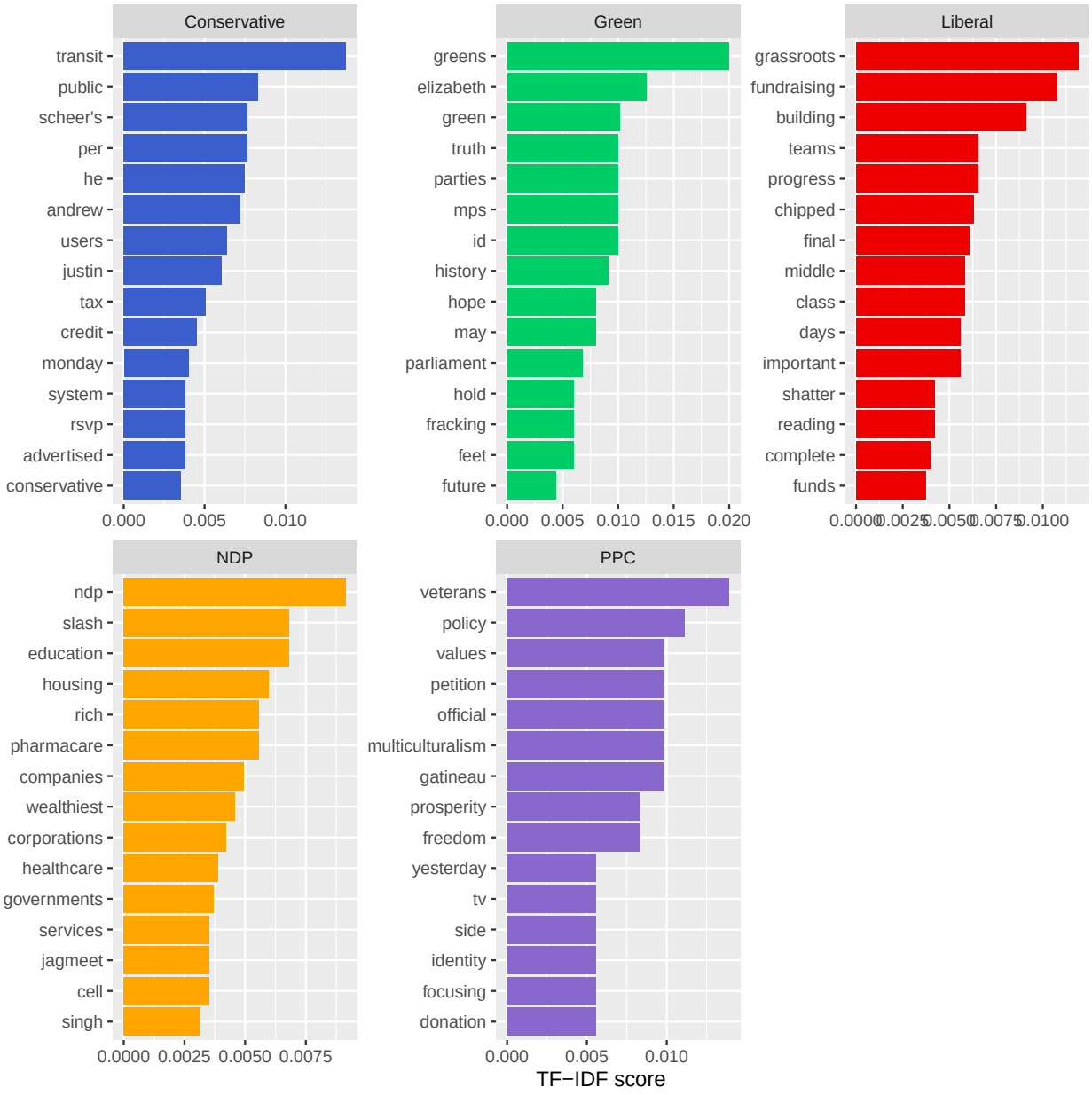


Figure 5: Top 15 extracted keywords of each party

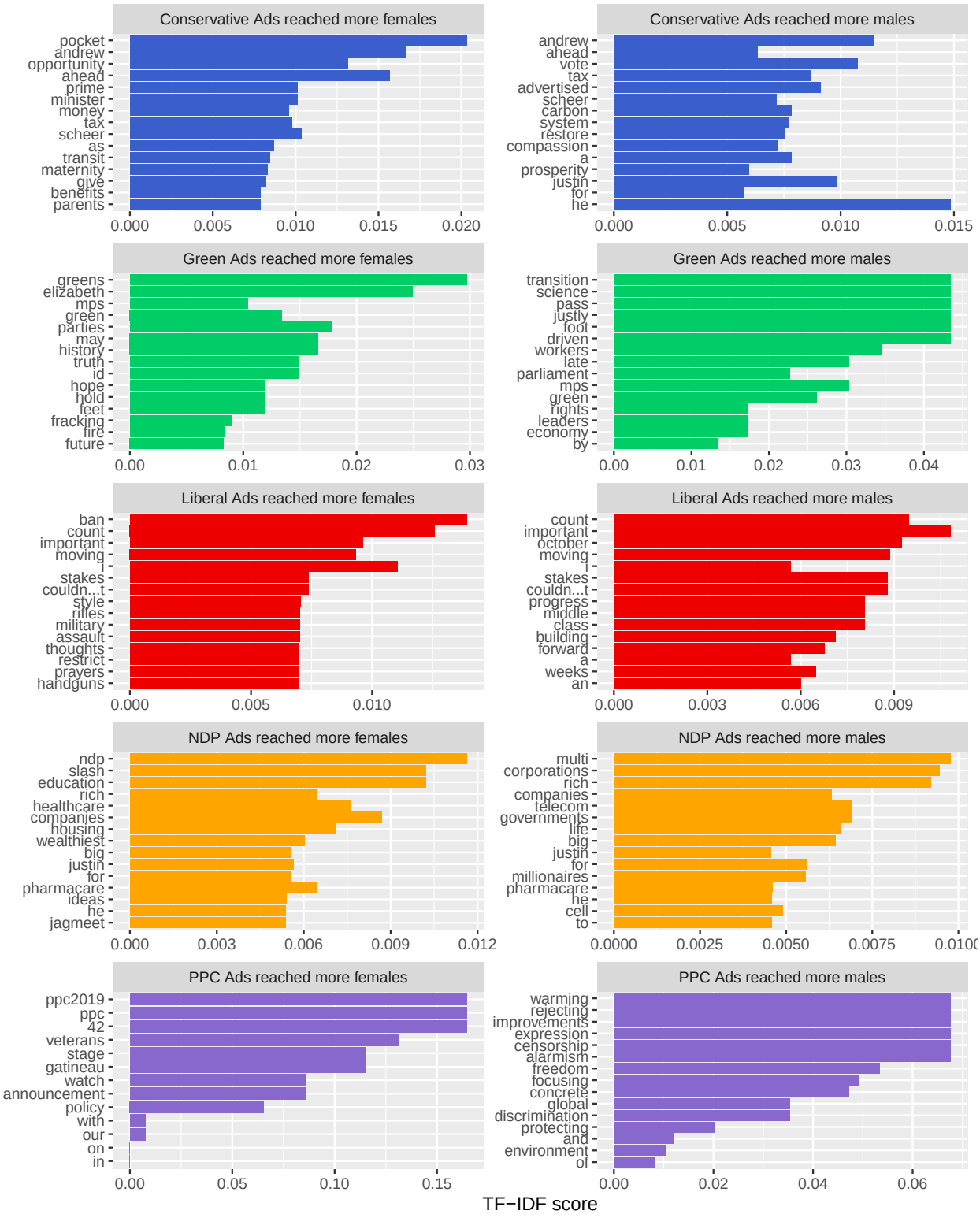


Figure 6: Top 15 extracted keywords of each party, segmented by gender

Election on Facebook Ads reflects the implicit limitation of women's participation in politics in Canada. Through the distribution landscape, we see that half of the parties are indirectly targeting a much larger amount of male audiences than female audiences. Through the different messaging, women (and men) more likely to be exposed to only certain topics and therefore polarize their social roles.

Appendix: All code for this report

```
knitr::opts_chunk$set(echo = TRUE)
library(tidyverse)
library(dplyr)
library(tidytext)
library(sentimentr)
library(cld3)
# Set so that long lines in R will be wrapped:
knitr::opts_chunk$set(tidy.opts=list(width.cutoff=80), echo = FALSE, warning = FALSE)
#### Loading all csv files ####
liberal_party_ads <- read.csv("Liberal_party/fb_ads.csv")
liberal_party_ads_demos <- read.csv("Liberal_party/fb_ads_demos.csv")
liberal_party_ads_regions <- read.csv("Liberal_party/fb_ads_regions.csv")
liberal_candid_ads <- read.csv("Liberal_candid/fb_ads.csv")
liberal_candid_ads_demos <- read.csv("Liberal_candid/fb_ads_demos.csv")
liberal_candid_ads_regions <- read.csv("Liberal_candid/fb_ads_regions.csv")

conserv_party_ads <- read.csv("Conservative_party/fb_ads.csv")
conserv_party_ads_demos <- read.csv("Conservative_party/fb_ads_demos.csv")
conserv_party_ads_regions <- read.csv("Conservative_party/fb_ads_regions.csv")
conserv_candid_ads <- read.csv("Conservative_candid/fb_ads.csv")
conserv_candid_ads_demos <- read.csv("Conservative_candid/fb_ads_demos.csv")
conserv_candid_ads_regions <- read.csv("Conservative_candid/fb_ads_regions.csv")

ndp_party_ads <- read.csv("NDP_party/fb_ads.csv")
ndp_party_ads_demos <- read.csv("NDP_party/fb_ads_demos.csv")
ndp_party_ads_regions <- read.csv("NDP_party/fb_ads_regions.csv")
ndp_candid_ads <- read.csv("NDP_candid/fb_ads.csv")
ndp_candid_ads_demos <- read.csv("NDP_candid/fb_ads_demos.csv")
ndp_candid_ads_regions <- read.csv("NDP_candid/fb_ads_regions.csv")

blocq_party_ads <- read.csv("BlocQ_party/fb_ads.csv")
blocq_party_ads_demos <- read.csv("BlocQ_party/fb_ads_demos.csv")
blocq_party_ads_regions <- read.csv("BlocQ_party/fb_ads_regions.csv")
blocq_candid_ads <- read.csv("BlocQ_candid/fb_ads.csv")
blocq_candid_ads_demos <- read.csv("BlocQ_candid/fb_ads_demos.csv")
blocq_candid_ads_regions <- read.csv("BlocQ_candid/fb_ads_regions.csv")

green_party_ads <- read.csv("Green_party/fb_ads.csv")
green_party_ads_demos <- read.csv("Green_party/fb_ads_demos.csv")
green_party_ads_regions <- read.csv("Green_party/fb_ads_regions.csv")
# No ads found in the Green Party candidate page

ppc_party_ads <- read.csv("PPC_party/fb_ads.csv")
ppc_party_ads_demos <- read.csv("PPC_party/fb_ads_demos.csv")
ppc_party_ads_regions <- read.csv("PPC_party/fb_ads_regions.csv")
# No ads found in the PPC candidate page
#### Set colours for graphs ####
cl_female = "#e96e61"
cl_male = "#535ac2"
cl_unknown = "darkgrey"
gender_color <- scale_fill_manual("legend",
```

```
                    values = c( "female" = cl_female,
                                "male"   = cl_male,
                                "unknown"= cl_unknown))
age_color <- scale_fill_manual("legend",
                    values = c("18-24" = "#e34434",
                               "25-34" = "#e96e61",
                               "35-44" = "#a9ace0",
                               "45-54" = "#7e83d1",
                               "55-64" = "#535ac2",
                               "65+"   = "#2931b3"))
party_color <- scale_fill_manual("legend",
                    values = c("Liberal" = "red2",
                               "Conservative" = "royalblue3",
                               "NDP" = "Orange",
                               "Bloc Québécois" = "lightblue",
                               "Green" ="springgreen3",
                               "PPC" = "mediumpurple3"))
#### Function to merge party and candidate data ####
merge_party_candid <- function(party_ads, candid_ads,
                               party_ads_demos, candid_ads_demos,
                               party_ads_regions, candid_ads_regions,
                               party_name) {

  ads <- rbind(party_ads, candid_ads) %>%
  mutate_at(c("ad_id", "page_id"), as.factor) %>%
  mutate_at(c("ad_delivery_start_time", "ad_delivery_stop_time"), as.Date) %>%
  mutate(ad_creative_body = gsub("http.*","",  ad_creative_body)) %>%
  mutate(party = party_name) %>%
  tibble()

  ads_demos <-
  rbind(party_ads_demos, candid_ads_demos) %>%
  mutate(ad_id = as.factor(ad_id)) %>%
  mutate(party = party_name) %>%
  tibble()

  ads_regions <-
  rbind(party_ads_regions, candid_ads_regions) %>%
  mutate(ad_id = as.factor(ad_id))%>%
  mutate(party = party_name) %>%
  tibble()

  return(list("ads" = ads,
              "ads_demos" = ads_demos,
              "ads_regions" = ads_regions))
}
#### Merge data for all parties ####
liberal <- merge_party_candid(party_ads = liberal_party_ads,
                   candid_ads = liberal_candid_ads,
                   party_ads_demos = liberal_party_ads_demos,
                   candid_ads_demos = liberal_candid_ads_demos,
                   party_ads_regions = liberal_party_ads_regions,
                   candid_ads_regions = liberal_candid_ads_regions,
```

```
                    party_name = "Liberal")

conserv <- merge_party_candid(party_ads = conserv_party_ads,
                    candid_ads = conserv_candid_ads,
                    party_ads_demos = conserv_party_ads_demos,
                    candid_ads_demos = conserv_candid_ads_demos,
                    party_ads_regions = conserv_party_ads_regions,
                    candid_ads_regions = conserv_candid_ads_regions,
                    party_name = "Conservative")

ndp <- merge_party_candid(party_ads = ndp_party_ads,
                    candid_ads = ndp_candid_ads,
                    party_ads_demos = ndp_party_ads_demos,
                    candid_ads_demos = ndp_candid_ads_demos,
                    party_ads_regions = ndp_party_ads_regions,
                    candid_ads_regions = ndp_candid_ads_regions,
                    party_name = "NDP")

blocq <- merge_party_candid(party_ads = blocq_party_ads,
                    candid_ads = blocq_candid_ads,
                    party_ads_demos = blocq_party_ads_demos,
                    candid_ads_demos = blocq_candid_ads_demos,
                    party_ads_regions = blocq_party_ads_regions,
                    candid_ads_regions = blocq_candid_ads_regions,
                    party_name = "Bloc Québécois")

green_party_ads <- mutate_at(green_party_ads, c("ad_id", "page_id"), as.factor) %>%
  mutate(party = "Green") %>%
  mutate(ad_creative_body = gsub("http.*","",  ad_creative_body)) %>% tibble()
green_party_ads_demos <- mutate(green_party_ads_demos, ad_id = as.factor(ad_id)) %>%
  mutate(party = "Green") %>% tibble()
green_party_ads_regions <- mutate(green_party_ads_regions, ad_id = as.factor(ad_id)) %>%
  mutate(party = "Green") %>% tibble()
ppc_party_ads <- mutate_at(ppc_party_ads, c("ad_id", "page_id"), as.factor) %>%
  mutate(party = "PPC") %>%
  mutate(ad_creative_body = gsub("http.*","",  ad_creative_body)) %>% tibble()
ppc_party_ads_demos <- mutate(ppc_party_ads_demos, ad_id = as.factor(ad_id)) %>%
  mutate(party = "PPC") %>% tibble()
ppc_party_ads_regions <- mutate(ppc_party_ads_regions, ad_id = as.factor(ad_id))%>%
  mutate(party = "PPC") %>% tibble()
## Impression and Spending by Regions
liberal_ads_imp_spend <-select(liberal$ads, impressions_min, spend_min, ad_id, party)
conserv_ads_imp_spend <-select(conserv$ads, impressions_min, spend_min, ad_id, party)
ndp_ads_imp_spend <-select(ndp$ads, impressions_min, spend_min, ad_id, party)
blocq_ads_imp_spend <-select(blocq$ads, impressions_min, spend_min, ad_id, party)
green_ads_imp_spend <-select(green_party_ads, impressions_min, spend_min, ad_id, party)
ppc_ads_imp_spend <-select(ppc_party_ads, impressions_min, spend_min, ad_id, party)

calc_by_region <- function(ad_regions, ads_imp_spend, party_name) {
  right_join(ad_regions, ads_imp_spend, by= "ad_id") %>%
  mutate(impression = impressions_min*percentage) %>%
  mutate(spend = spend_min*percentage) %>%
  group_by(region) %>%
```

```
  summarise_at(c("impression","spend"), sum) %>%
  mutate(party = party_name) %>%
  mutate(impression_pt = round(impression/sum(impression),4)*100) %>%
  mutate(spend_pt = round(spend/sum(spend), 4)*100)
}

liberal_region <- calc_by_region(liberal$ads_regions,
                                 liberal_ads_imp_spend,liberal$ads$party[1])
conserv_region <- calc_by_region(conserv$ads_regions,
                                 conserv_ads_imp_spend,conserv$ads$party[1])
ndp_region <- calc_by_region(ndp$ads_regions,
                             ndp_ads_imp_spend,ndp$ads$party[1])
blocq_region <- calc_by_region(blocq$ads_regions,
                               blocq_ads_imp_spend,blocq$ads$party[1])
green_region <- calc_by_region(green_party_ads_regions,
                               green_ads_imp_spend,green_party_ads$party[1])
ppc_region <- calc_by_region(ppc_party_ads_regions,
                             ppc_ads_imp_spend,ppc_party_ads$party[1])
all_ads_summary <- rbind(liberal$ads, conserv$ads, ndp$ads,
                         blocq$ads, green_party_ads, ppc_party_ads) %>%
  mutate(ad_count_total = 1)%>%
  group_by(party) %>%
  summarise_at(c("ad_count_total", "spend_min", "spend_max",
                 "impressions_min", "impressions_max"), sum, na.rm = TRUE)

colnames(all_ads_summary) <- c("Party", "Ad Counts", "Min Spending",
                       "Max Spending", "Min Impressions", "Max Impressions")


knitr::kable(all_ads_summary[order(-all_ads_summary["Min Impressions"]), ],
booktabs = T, caption = "Ad counts, spending and impressions of each party")
all_regions_summary <- rbind(liberal_region, conserv_region, ndp_region,
                             blocq_region, green_region, ppc_region)

ggplot(all_regions_summary, aes(y=reorder(region, spend_pt), x=spend_pt)) +
        geom_bar(stat = "identity", aes(fill = party)) +
        geom_text(aes(label = paste(spend_pt, "%")), size=3,
                  hjust=ifelse(all_regions_summary$spend_pt>80, 1, 0))+
        facet_wrap(. ~ party, scales = "free_y", nrow=2) +
        theme(legend.position = "none") +
        party_color +
        labs(x="", y="")
calc_by_demos <- function(ad_demos, ads_imp_spend, party_name) {
  right_join(ad_demos, ads_imp_spend, by= "ad_id") %>%
  mutate(impression = impressions_min*percentage) %>%
  mutate(spend = spend_min*percentage) %>%
  group_by(age, gender) %>%
  summarise_at(c("impression","spend"), sum, na.rm =T) %>%
  mutate(party = party_name) %>%
  ungroup() %>%
  mutate(impression_pt = round(impression/sum(impression),4)*100) %>%
  mutate(spend_pt = round(spend/sum(spend), 4)*100)
}
```

```
liberal_demos <- calc_by_demos(liberal$ads_demos,
                               liberal_ads_imp_spend,liberal$ads$party[1])
conserv_demos <- calc_by_demos(conserv$ads_demos,
                               conserv_ads_imp_spend,conserv$ads$party[1])
ndp_demos <- calc_by_demos(ndp$ads_demos,
                           ndp_ads_imp_spend,ndp$ads$party[1])
blocq_demos <- calc_by_demos(blocq$ads_demos,
                             blocq_ads_imp_spend,blocq$ads$party[1])
green_demos <- calc_by_demos(green_party_ads_demos,
                             green_ads_imp_spend,green_party_ads$party[1])
ppc_demos <- calc_by_demos(ppc_party_ads_demos,
                           ppc_ads_imp_spend,ppc_party_ads$party[1])
all_demos_summary <- rbind(liberal_demos, conserv_demos,
                           ndp_demos, blocq_demos, green_demos, ppc_demos)
gender_demos_summary <- all_demos_summary %>%
  group_by(gender, party) %>%
  summarise(gender_impression_pt = sum(impression_pt), .groups = "drop") %>%
  mutate(party = fct_relevel(party, "Green", "Liberal",
                             "NDP", "Conservative", "Bloc Québécois", "PPC"))
  ggplot(gender_demos_summary, aes(fill=gender, x=gender_impression_pt, y=party)) +
  geom_bar(position="fill", stat="identity") +
  geom_text(position = "fill", aes(label =
      paste(gender_impression_pt, "%")), size=3, colour = "Beige",
      hjust= ifelse(gender_demos_summary$gender_impression_pt<5, 0.25, 1), vjust = 0)+
  gender_color+labs(x="Spending percentage by gender")
all_ads_text <- rbind(liberal$ads, conserv$ads, ndp$ads,
                      blocq$ads, green_party_ads, ppc_party_ads) %>%
  select(ad_id, ad_creative_body) %>%
  mutate(language = cld3::detect_language(ad_creative_body))%>%
  mutate(sentiment_score =
           sentimentr::sentiment_by(ad_creative_body)$ave_sentiment) %>%
  mutate(sentiment =
           ifelse(sign(sentiment_score) >=0 , "positive", "negative"))

all_ads_demos <- rbind(liberal$ads_demos, conserv$ads_demos,
  ndp$ads_demos, green_party_ads_demos, blocq$ads_demos, ppc_party_ads_demos) %>%
  select(-age) %>%
  group_by(ad_id,gender ) %>%
  mutate(percentage = sum(percentage)) %>%
  ungroup() %>%
  unique() %>%
  pivot_wider(names_from = gender, values_from = percentage) %>%
  mutate(dom_gender = ifelse(sign(male-female) > 0,
        "Ads reached more males", "Ads reached more females")) %>%
  select(ad_id, party, dom_gender) %>%
  left_join(all_ads_text, by="ad_id")

table <- table(all_ads_demos$dom_gender, all_ads_demos$party)
prop_table <- prop.table(table(all_ads_demos$dom_gender,
              all_ads_demos$party), 2) %>% round(2)
knitr::kable(table, booktabs = T,
caption = "Number of more-female-reaching ads and more-male-reaching ads")
knitr::kable(prop_table, booktabs = T,
```

```
caption = "Proprotions of more-female-reaching ads and more-male-reaching ads")
age_demos_summary <- all_demos_summary %>%
  group_by(age, party) %>%
  summarise(age_impression_pt = sum(impression_pt), .groups = "drop") %>%
  mutate(party = fct_relevel(party, "NDP", "Green",  "PPC",
                             "Liberal", "Bloc Québécois", "Conservative" ))
  age_demos_summary_graph <- filter(age_demos_summary, age != "13-17")
  ggplot(age_demos_summary_graph, aes(fill=age,
                                      x=age_impression_pt, y=party)) +
  geom_bar(position="fill", stat="identity") +
  geom_text(position = "fill", aes(label = paste(age, "\n",
              age_impression_pt, "%")), colour = "Beige",
              size=3, hjust= 1, vjust = .5)+
    age_color+ labs(x="Spending percentage by age group")
unique_ads <- rbind(liberal$ads, conserv$ads, ndp$ads,
                    blocq$ads, green_party_ads, ppc_party_ads) %>%
  select("ad_creative_body", "party") %>%
  unique() %>%
  mutate(language = cld3::detect_language(ad_creative_body))%>%
  mutate(language = str_replace_all(language, "ja", "zh")) %>%
  mutate(language = str_replace_all(language, "cy", "en")) %>%
  mutate(language = str_replace_all(language, "fil", "en"))

en_unique_ads <- filter(unique_ads, language == "en") %>%
  mutate(sentiment_score =
           sentimentr::sentiment_by(ad_creative_body)$ave_sentiment) %>%
  mutate(sentiment = ifelse(sign(sentiment_score) >=0 , "positive", "negative"))

en_unique_ads_summary <- en_unique_ads %>%
  group_by(party, sentiment) %>%
  mutate(count = n()) %>%
  select(party, sentiment, count) %>%
  unique() %>%
  ungroup()%>%
  group_by(party) %>%
  mutate(sentiment_pt = round(count/sum(count),4)*100)

yarrr::pirateplot(formula =  sentiment_score ~ party,
   data = en_unique_ads,
   xlab = NULL, ylab = "Sentiment Score",
   main = "Sentiment Score per Party",
   pal = c("royalblue3", "springgreen3", "red2", "Orange", "mediumpurple3"),
   point.o = .2,
   avg.line.o = 1,
   theme = 0,
   point.pch = 16,
   point.cex = 1.5,
   jitter.val = .1,
   cex.lab = .9, cex.names = .7)
library(png)
library(grid)
img <- readPNG("images/conserv_neg_ad.png")
grid.raster(img)
```

```
en_unique_ads_words <- en_unique_ads %>%
  tidytext::unnest_tokens(word, ad_creative_body) %>%
  count(party, word, sort = TRUE)

total_words <- en_unique_ads_words %>%
  group_by(party) %>% summarize(total = sum(n)) %>% ungroup()
en_unique_ads_words <-
  left_join(en_unique_ads_words, total_words, by = "party")

en_unique_ads_words_tfidf <- en_unique_ads_words %>%
 tidytext::bind_tf_idf(word, party, n)

en_unique_ads_words_tfidf <- en_unique_ads_words_tfidf %>%
  select(-total) %>%
  arrange(party, desc(tf_idf))

en_unique_ads_words_tfidf %>%
  group_by(party) %>%
  slice_max(tf_idf, n = 15) %>%
  ungroup() %>%
  ggplot(aes(tf_idf, fct_reorder(word, tf_idf), fill = party)) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~party, ncol = 3, scales = "free") +
  labs(x = "tf-idf", y = NULL) +
  party_color+
  labs(x="TF-IDF score")
en_unique_ads_words <- all_ads_demos %>%
  filter(language == "en") %>%
  mutate(party_dom_gender = str_c(party, dom_gender, sep = " "))%>%
  tidytext::unnest_tokens(word, ad_creative_body) %>%
  count(party_dom_gender, party, word, sort = TRUE)

total_words <- en_unique_ads_words %>% group_by(party_dom_gender) %>%
  summarize(total = sum(n)) %>% ungroup()
en_unique_ads_words <- left_join(en_unique_ads_words, total_words,
                                 by = "party_dom_gender")

en_unique_ads_words_tfidf <- en_unique_ads_words %>%
 tidytext::bind_tf_idf(word, party_dom_gender, n)

en_unique_ads_words_tfidf <- en_unique_ads_words_tfidf %>%
  select(-total) %>%
  arrange(party_dom_gender, desc(tf_idf))

en_unique_ads_words_tfidf %>%
  group_by(party_dom_gender) %>%
  slice_max(tf_idf, n = 15) %>%
  ungroup() %>%
  ggplot(aes(tf_idf, fct_reorder(word, tf_idf), fill = party)) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~party_dom_gender, ncol = 2, scales = "free") +
  labs(x = "tf-idf", y = NULL)+party_color+
  labs(x="TF-IDF score")
```

References

Silva, Marcio, Lucas Santos de Oliveira, Athanasios Andreou, Pedro Olmo Vaz de Melo, Oana Goga, and Fabricio Benevenuto. 2020. "Facebook Ads Monitor: An Independent Auditing System for Political Ads on Facebook." In *Proceedings of the Web Conference 2020*, 224–34.

"Statista." 2021. *Statista*. https://www.statista.com/.

The Guardian, The Guardian. 2019. *Canada Election 2019: Full Results*, October.

Woolf, Max. n.d. *Facebook-Ad-Library-Scraper*. Github. https://github.com/minimaxir/facebook-ad-library-scraper.