**Codes used to generate the tables and analyses analyzed in the research**

**Table of Content**

[Table 1: 2](#_Toc208609492)

[Table 2: 2](#_Toc208609493)

[Table 3: 2](#_Toc208609494)

[Table 4: library(dplyr) library(knitr) 3](#_Toc208609495)

[Table 5: 3](#_Toc208609496)

[Table: 6 3](#_Toc208609497)

[Table 7: 4](#_Toc208609498)

[Table 8: 5](#_Toc208609499)

[Table 9: 6](#_Toc208609500)

[Table 10. Party preference by gender 6](#_Toc208609501)

[Table 11: Distribution of voters’ preference towards spending on education by gender 7](#_Toc208609502)

[Table 12. : Distribution of voters’ preference towards spending on affordable housing by gender 8](#_Toc208609503)

[Table 13: Regression with stargazer - excluding gender as controlling variable 8](#_Toc208609504)

[Table 14. Stargazer on interaction (1) and Base (2) models 9](#_Toc208609505)

[Test statistics: Tables 15-22 9](#_Toc208609506)

[Table 23: Multilevel (hierarchical) logistic regression model 10](#_Toc208609507)

[Table 24. Stargazer including the linear mixed models 1 and 2: 10](#_Toc208609508)

[Table 25. Multinominal regression model: 10](#_Toc208609509)

## Table 1:

library(dplyr) library(knitr)

voting\_counts <- c( "Liberal Party" = 890,

"Conservative Party" = 3637, "NDP" = 2828

)

voting\_data <- as.data.frame(voting\_counts)

colnames(voting\_data) <- c("Frequency") voting\_data$Category <- rownames(voting\_data) rownames(voting\_data) <- NULL

voting\_data <- voting\_data %>%

mutate(Percent = round(Frequency / sum(Frequency) \* 100, 2), Cumulative\_Percent = round(cumsum(Percent), 2))

voting\_data <- voting\_data %>%

select(Category, Frequency, Percent, Cumulative\_Percent)

print(kable(voting\_data, caption = "Frequency Table of Voting Parties"))

## Table 2:

library(dplyr) library(knitr)

cps21\_spend\_data <- data.frame(

Category = c("Spending Less", "Spending the Same", "Spending More"),

Frequency = c(633, 2562, 6925)

)

cps21\_spend\_data <- cps21\_spend\_data %>%

mutate(Percent = round(Frequency / sum(Frequency) \* 100, 2), Cumulative\_Percent = round(cumsum(Percent), 2))

print(kable(cps21\_spend\_data, caption = "Frequency Table for Affordable Housing Spending Responses"))

## Table 3:

library(dplyr) library(knitr)

categories <- c("Spending Less", "Spending the Same", "Spending More") frequencies <- c(823, 7176, 12200)

cps21\_data <- data.frame(Category = categories, Frequency = frequencies)

cps21\_data <- cps21\_data %>%

mutate(Percent = round(Frequency / sum(Frequency) \* 100, 2), Cumulative\_Percent = round(cumsum(Percent), 2))

print(kable(cps21\_data, caption = "Frequency Table for Spend on Education"))

## Table 4: library(dplyr) library(knitr)

gender\_data <- data.frame(

Gender = c("Man", "Woman"), Frequency = c(9474, 11370)

)

gender\_data <- gender\_data %>%

mutate(Percent = round(Frequency / sum(Frequency) \* 100, 2), Cumulative\_Percent = round(cumsum(Percent), 2))

gender\_data <- gender\_data %>%

select(Gender, Frequency, Percent, Cumulative\_Percent)

print(kable(gender\_data, caption = "Frequency Table for Gender"))

## Table 5:

summary\_stats <- summary(mydata11$cps21\_spend\_educ) sd\_educ <- sd(mydata11$cps21\_spend\_educ, na.rm = TRUE) summary\_table <- data.frame(

Statistic = c("Min", "1st Qu.", "Median", "Mean", "3rd Qu.", "Max", "Standard Deviation"), Value = c(summary\_stats, sd\_educ)

)

print(summary\_table) library(knitr)

kable(summary\_table, caption = "Summary Statistics for Education Spending") library(knitr)

formatted\_table <- data.frame(

Statistic = c("Min", "1st Qu.", "Median", "Mean", "3rd Qu.", "Max", "Standard Deviation"), Value = round(c(summary\_stats, sd\_educ), 2)

)

kable(formatted\_table, caption = "Summary Statistics for Education Spending")

## Table: 6

summary\_stats <- summary(mydata11$cps21\_spend\_afford\_h)

sd\_afford\_h <- round(sd(mydata11$cps21\_spend\_afford\_h, na.rm = TRUE), 2) summary\_table <- data.frame(

Statistic = c("Minimum", "1st Quartile", "Median", "Mean", "3rd Quartile", "Maximum", "Standard Deviation"),

Value = round(c(summary\_stats[1], summary\_stats[2], summary\_stats[3], summary\_stats[4], summary\_stats[5], summary\_stats[6], sd\_afford\_h), 2)

)

print(summary\_table) library(knitr)

kable(summary\_table, caption = "Summary Statistics for Housing Spending")

Boxplot of education spending:

boxplot(mydata11$cps21\_spend\_educ, main = "Boxplot of Education Spending", ylab = "Spending")

Histogram of housing spending:

hist(mydata11$cps21\_spend\_afford\_h, main = "Histogram of Housing Spending", xlab = "Spending", breaks = 10)

## Table 7:

sd\_votechoice <- sd(mydata11$cps21\_votechoice, na.rm = TRUE) cat("Standard Deviation for Vote Choice: ", round(sd\_votechoice, 2), "\n") summary\_stats\_votechoice <- summary(mydata11$cps21\_votechoice) summary\_table\_votechoice <- data.frame(

Statistic = c("Minimum", "1st Quartile", "Median", "Mean", "3rd Quartile", "Maximum", "Standard Deviation"),

Value = round(c( summary\_stats\_votechoice[1], summary\_stats\_votechoice[2], summary\_stats\_votechoice[3], summary\_stats\_votechoice[4], summary\_stats\_votechoice[5],

summary\_stats\_votechoice[6], round(sd\_votechoice, 2)

), 2)

)

print(summary\_table\_votechoice) install.packages("kableExtra") library(kableExtra) summary\_table\_votechoice <- data.frame(

Statistic = c("Minimum", "1st Quartile", "Median", "Mean", "3rd Quartile", "Maximum", "Standard Deviation"),

Value = round(c(1.00, 1.00, 2.00, 3.03, 4.00, 7.00, 2.02), 2)

)

summary\_table\_votechoice %>% kable() %>%

kable\_styling(full\_width = FALSE, position = "left") kable(summary\_table\_votechoice, caption = "Summary Statistics for Voting Choice")

library(dplyr) library(ggplot2) library(haven)

mydata$pes21\_rural\_urban <- as\_factor(mydata$pes21\_rural\_urban)

Distribution of Areas of living / barchart mydata <- mydata %>% mutate(living\_area\_group = case\_when(

pes21\_rural\_urban %in% c("1", "2", "3") ~ "Rural", # Group 1, 2, 3 as Rural pes21\_rural\_urban == "4" ~ "Suburb", # Group 4 as Suburb pes21\_rural\_urban == "5" ~ "Urban", # Group 5 as Urban

pes21\_rural\_urban == "6" ~ "Don't Know", # Group 6 as Don't Know TRUE ~ "Other" # Any other category as Other (optional)

))

table(mydata$living\_area\_group)

ggplot(mydata, aes(x = living\_area\_group, fill = living\_area\_group)) + geom\_bar() + # Create bars

scale\_fill\_brewer(palette = "Set3") + # Specify color palette labs(title = "Distribution of Areas of Living",

x = "Areas of Living", y = "Count",

fill = "Living Area") +

theme\_minimal() + # Apply minimal theme

theme(axis.text.x = element\_text(angle = 45, hjust = 1)) # Rotate x-axis text for better readability

Born in Canada /Frequency of Immigration status

library(dplyr) library(knitr)

mydata <- data.frame(

cps21\_bornin\_canada = c(rep(1, 8), rep(2, 3), rep(3, 1)) # 8 ones, 3 twos, and 1 three

)

summary\_stats <- mydata %>% summarise(

Minimum = min(cps21\_bornin\_canada, na.rm = TRUE),

Q1 = quantile(cps21\_bornin\_canada, 0.25, na.rm = TRUE), # First quartile Median = median(cps21\_bornin\_canada, na.rm = TRUE),

Mean = mean(cps21\_bornin\_canada, na.rm = TRUE),

Q3 = quantile(cps21\_bornin\_canada, 0.75, na.rm = TRUE), # Third quartile Maximum = max(cps21\_bornin\_canada, na.rm = TRUE),

Standard\_Deviation = sd(cps21\_bornin\_canada, na.rm = TRUE)

)

print(kable(summary\_stats, caption = "Summary Statistics for Immigration Status (Born in Canada)"))

Vote choice distribution

barplot(vote\_counts, main = "Vote Choice Distribution", xlab = "Vote Choice", ylab = "Frequency")

## Table 8:

cont\_table1 <- table(mydata$cps21\_genderid, mydata$vcps21\_votechoice, mydata$cps21\_spend\_educ)

cont\_df <- as.data.frame(cont\_table1)

colnames(cont\_df) <- c("Gender", "VoteChoice", "SpendEducation", "Frequency") cont\_df <- cont\_df %>%

group\_by(Gender, VoteChoice) %>%

mutate(Percent = Frequency / sum(Frequency) \* 100, Cumulative\_Freq = cumsum(Frequency), Cumulative\_Percent = cumsum(Percent))

print(kable(cont\_df, caption = "Cumulative Association Table for Gender, Voting Choice, and Spending on Education"))

## Table 9:

mydata <- mydata %>% mutate(

Gender = as.numeric(factor(cps21\_genderid, levels = c("Man", "Woman"))),

VoteChoice = as.numeric(factor(vcps21\_votechoice, levels = c("Liberal Party", "Conservative Party", "NDP"))),

SpendHouse = as.numeric(factor(cps21\_spend\_afford\_h, levels = c("Less", "Same", "More")))

)

print(head(mydata))

cont\_table <- table(mydata$Gender, mydata$VoteChoice, mydata$SpendHouse) cont\_df <- as.data.frame(cont\_table)

colnames(cont\_df) <- c("Gender", "VoteChoice", "SpendHouse", "Frequency") cont\_df <- cont\_df %>%

group\_by(Gender, VoteChoice) %>% mutate(

Percent = Frequency / sum(Frequency) \* 100, Cumulative\_Freq = cumsum(Frequency), Cumulative\_Percent = round(cumsum(Percent), 2)

)

cont\_df <- cont\_df %>% mutate(

Gender = factor(Gender, levels = c(1, 2), labels = c("Man", "Woman")),

VoteChoice = factor(VoteChoice, levels = c(1, 2, 3), labels = c("Liberal Party", "Conservative Party", "NDP")),

SpendHouse = factor(SpendHouse, levels = c(1, 2, 3), labels = c("Less", "Same", "More"))

)

print(kable(cont\_df, caption = "Cumulative Association Table with Categorical Associations for Gender, Voting Choice, and Spending"))

## Table 10. Party preference by gender

# Install haven necessary to read dataset install.packages("haven") #Load haven library library("haven")

data <- X2021\_Canadian\_Election\_Study\_v2\_0

write.csv(data, "C:\\Users\\arnol\\Desktop\\Writing\\2021 Canadian Election Study v2.0 (1).csv", row.names = FALSE)

data\_csv <- read.csv("C:\\Users\\arnol\\Desktop\\Writing\\2021 Canadian Election Study v2.0 (1).csv") install.packages("ggplot2")

library(ggplot2)

data$cps21\_votechoice <- factor(data$cps21\_votechoice) data$cps21\_genderid <- factor(data$cps21\_genderid) library(ggplot2)

# Ensure that cps21\_votechoice and cps21\_genderid are treated as factors data$cps21\_votechoice <- as.factor(data$cps21\_votechoice) data$cps21\_genderid <-

as.factor(data$cps21\_genderid) # Filter for only parties 1, 2, and 3

my\_data\_filtered <- data[data$cps21\_votechoice %in% c("1", "2", "3"), ] # Filter for only Men and Women

my\_data\_filtered <- my\_data\_filtered[my\_data\_filtered$cps21\_genderid %in% c("1", "2"), ]

# Define labels for party and gender

party\_labels <- c("1" = "Liberal", "2" = "Conservative", "3" = "NDP") gender\_labels <- c("1" = "Men", "2" = "Women")

# Define custom colors

custom\_colors <- c("1" = "#ADD8E6", "2" = "#F08080") # Light Blue for Men, Light Coral for Women # Grouped Bar Chart (Side-by-Side, with parties on the x-axis)

ggplot(my\_data\_filtered, aes(x = cps21\_votechoice, fill = cps21\_genderid)) + geom\_bar(position = "dodge") + # "dodge" for side-by-side

labs(x = "Party Preference", y = "Count", title = "Party Preference by Gender (Men and Women only)") + scale\_fill\_manual(values = custom\_colors, name = "Gender", labels = gender\_labels) + # Changed here scale\_x\_discrete(labels = party\_labels) +

theme\_bw()

## Table 11: Distribution of voters’ preference towards spending on education by gender

library(ggplot2) library(dplyr)

library(haven) # Make sure you have this installed too data <- X2021\_Canadian\_Election\_Study\_v2\_0 # Ensuring cps21\_votechoice and cps21\_genderid are treated as factors data$cps21\_votechoice <-

as.factor(data$cps21\_votechoice) data$cps21\_genderid <- as.factor(data$cps21\_genderid)

# Convert haven labelled to factor and then numeric if needed

data$cps21\_spend\_educ <- as\_factor(data$cps21\_spend\_educ) # Convert labelled to factors # if you need the underlying numeric representation

# my\_data$cps21\_spend\_educ <- as.numeric(my\_data$cps21\_spend\_educ)

# Filter for only parties 1, 2, and 3 data\_filtered <- data %>% filter(cps21\_votechoice %in% c("1", "2", "3"))

# Filter for only Men and Women data\_filtered <- data\_filtered %>% filter(cps21\_genderid %in% c("1", "2")) # Define labels for party and gender

party\_labels <- c("1" = "Liberal", "2" = "Conservative", "3" = "NDP") gender\_labels <- c("1" = "Men", "2" = "Women")

# Histogram with faceting

ggplot(data\_filtered, aes(x = cps21\_spend\_educ, fill = cps21\_genderid)) + geom\_bar(position = "dodge") + # Changed to geom\_bar to display categories

facet\_wrap(~ factor(cps21\_votechoice, levels = c("1", "2", "3"), labels = party\_labels), ncol = 1) + #

Use party\_labels here

labs(x = "Spending on Education", y = "Count",

title = "Distribution of Spending on Education by Gender and Party Preference") + scale\_fill\_manual(values = c("1" = "#1b909a", "2" = "#7900f1"), # Custom colors

name = "Gender", labels = gender\_labels) + theme\_bw() +

guides(fill = guide\_legend(title = "Gender")) # Adding gender label to the legend

## Table 12. : Distribution of voters’ preference towards spending on affordable housing by gender

library(ggplot2) library(dplyr)

library(haven) # Make sure you have this installed too

data <- X2021\_Canadian\_Election\_Study\_v2\_0

# Ensuring cps21\_votechoice and cps21\_genderid are treated as factors data$cps21\_votechoice <- as.factor(data$cps21\_votechoice) data$cps21\_genderid <- as.factor(data$cps21\_genderid) data$cps21\_spend\_afford\_h <- as.factor(data$cps21\_spend\_afford\_h)

# Filter for only parties 1, 2, and 3 data\_filtered <- data %>% filter(cps21\_votechoice %in% c("1", "2", "3"))

# Filter for only Men and Women data\_filtered <- data\_filtered %>% filter(cps21\_genderid %in% c("1", "2"))

# Define labels for party and gender

party\_labels <- c("1" = "Liberal", "2" = "Conservative", "3" = "NDP") gender\_labels <- c("1" = "Men", "2" = "Women")

# Histogram with faceting

ggplot(data\_filtered, aes(x = factor(cps21\_spend\_afford\_h, levels = c("1", "2", "3", "4"), labels = housing\_labels), fill = cps21\_genderid)) +

geom\_bar(position = "dodge") + # Changed to geom\_bar to display categories

facet\_wrap(~ factor(cps21\_votechoice, levels = c("1", "2", "3"), labels = party\_labels), ncol = 1) + #

Use party\_labels here

labs(x = "Spending on Affordable Housing", y = "Count",

title = "Distribution of Spending on Affordable Housing by Gender and Party Preference") + scale\_fill\_manual(values = c("1" = "#8724f5", "2" = "#F0ad2f"), # Custom colors

name = "Gender", labels = gender\_labels) + theme\_bw() +

guides(fill = guide\_legend(title = "Gender")) # Adding gender label to the legend

## Table 13: Regression with stargazer - excluding gender as controlling variable

> mymodel1=lm(cps21\_votechoice ~ cps21\_genderid2 + cps21\_spend\_educ + cps21\_spend\_afford\_h, data=mydata1\_filtered)

> summary(mymodel1)

>mymodel2=lm(cps21\_votechoice ~ cps21\_genderid2 + cps21\_spend\_educ + pes21\_rural\_urban + cps21\_spend\_afford\_h + cps21\_employment + cps21\_education, data=mydata1\_filtered)

> summary(mymodel2)

>mymodel3=lm(cps21\_votechoice ~ cps21\_genderid2 + cps21\_spend\_educ + pes21\_rural\_urban + cps21\_employment + cps21\_education, data=mydata1\_filtered)

> summary(mymodel3)

>mymodel4=lm(cps21\_votechoice ~ cps21\_spend\_educ + cps21\_spend\_afford\_h, data=mydata1\_filtered)

> summary(mymodel4)

>mymodel5=lm(cps21\_votechoice ~ cps21\_spend\_educ + pes21\_rural\_urban + cps21\_spend\_afford\_h + cps21\_employment, data=mydata1\_filtered)

> summary(mymodel5) install.packages("stargazer") library(stargazer)

stargazer(mymodel1, mymodel2, mymodel3, mymodel4, mymodel5, type="text")

## Table 14. Stargazer on interaction (1) and Base (2) models

1. mydata$female <- ifelse(mydata$female == "Female", 1, 0)

mydata$spend\_gender\_interaction <- mydata$cps21\_spend\_educ \* mydata$female

mymodel\_interaction <- lm(cps21\_votechoice ~ cps21\_spend\_educ + female + spend\_gender\_interaction +

cps21\_spend\_afford\_h + cps21\_education + cps21\_employment + pes21\_rural\_urban + cps21\_bornin\_canada,

data = mydata)

summary(mymodel\_interaction)

2. base\_model <- lm(cps21\_votechoice ~ cps21\_spend\_educ + female + cps21\_spend\_afford\_h

+ cps21\_education + cps21\_employment + pes21\_rural\_urban + cps21\_bornin\_canada, data

= mydata) summary(base\_model) library(stargazer)

3. stargazer(mymodel\_interaction, base\_model, type="text")

## Test statistics: Tables 15-22

sigma(mydata1) summary(mydata1)$r.squared residuals( mydata1)

plot(fitted(mydata1), residuals(mydata1)) qqnorm(residuals(mydata1)) qqline(residuals(mydata1)) plot(density(res)) plot(density(residuals(mydata1))) plot(mydata1, which = 1)

library(lmtest) bptest(mydata1)

qqnorm(residuals(mydata2)) qqline(residuals(mydata2)) plot(density(res))

plot(density(residuals(mydata2)))

plot(mydata2, which = 1)

library(lmtest) bptest(mydata2)

## Table 23: Multilevel (hierarchical) logistic regression model

mydataa <- lmer(cps21\_votechoice ~ female + cps21\_spend\_educ + (1 | cps21\_province), data = mydata)

mydatab <- lmer(cps21\_votechoice ~ female + cps21\_spend\_educ + (1 + female | cps21\_province), data = mydata)

tab\_model(mydataa, mydatab)

Graph 1:

plot\_model(mydataa)

## Table 24. Stargazer including the linear mixed models 1 and 2:

stargazer(mydataa, mydatab, type="text")

## Table 25. Multinominal regression model:

mydata$cps21\_spend\_educC <- relevel(as.factor(mydata$cps21\_spend\_educC), ref = "Spend about the same as now")

mydata60<- multinom(cps21\_spend\_educC ~ gender\_factor2, data = mydata) summary(mydata60)

install.packages("sjPlot") library(sjPlot) tab\_model(mydata60)