

**Statistical approaches for finding the immigrant women's political participation in 2015 local elections and 2017 national elections**

**Using logistic regression, principle components analysis and principle components logistic regression**



**Amany El Gharib**

**Supervisor: Ingelin Steinsland**

**Department of Mathematics University of Bergen**

**Master of Science in Statistics (Data Analysis) STAT399**

**November 19, 2019**



# Summary

In this master thesis we study immigrant women's political participation in Norway during 2015 local elections and 2017 national elections. The data were collected through an online questionnaire. The constructing and building of this questionnaire were part of this thesis. It was built based on previous questionnaire at the same field. 95 women completed the questionnaire.

The data were analysed using; binomial multiple logistic regression, principle components analysis and principle components logistic regression.

The logistic regression and the principle components logistic regression results are very consistent and give almost the same findings. Furthermore the principle components logistic regression has added a bit more to the results of the logistic regression.

The main results show that the probability of voting in 2017 elections increases with older age, higher education, longer residency in Norway, stronger residency permit type. Furthermore the probability increases for people who have stronger reasons for voting, weaker reasons for not voting, higher participation in signing a petition and had a lower participation in the elections at the home country. It increase as well with higher participation; in a demonstration, in volunteering time not for political party, member of an interest group, as a member or a candidate of a political party and more interested in politics. Moreover it increases with stronger agree that "the Norwegian government does not care much about what people like them "immigrant women" think", "voting is a way to make their voices heard on issues they care about", "politician should focus more on Immigrant women's needs on their agendas" and "immigrant women SHOULD have more influence in politics in Norway". These people are financially independent.

From the other hand, the probability of voting in 2015 local elections increases with older age, higher education, moreover longer residency and higher Norwegian language level. It increases with lower participation in the elections at the home country , stronger reasons to vote, more interested in political participation, higher participation in signing a petition and weaker reasons to not vote. Furthermore, it increases with less degree of pressure on the participant's political opinion, more often the family used to discuss politics at the past and having less immigrant friends. In addition, the probability of voting in 2015 local elections increases for people who stronger agree with that "voting is a way to make the voice heard on issues they care about",

politicians don't put Immigrant women's needs at their agendas", "immigrant women SHOULD have more influence in politics in Norway" and they are financially independent.

It worth confirming that the probability of voting decreases for people who more agree with "Sometimes politics and government seem so complicated that a person like you cannot really understand what is going on" and "Traditions affect your political participation"

DEDICATED TO MY BELOVED  
PARENTS,

MY LOVELY HUSBAND AND

LITTLE ANGEL JOSEF

# Preface

This thesis is the final work of my study at the data analysis master program at the university of Berge. Since I do believe in women's political participation in general and the immigrant women's political participation in Norway in specific, the study in this thesis is about immigrant women's political participation in Norway in 2015 local elections and 2017 national elections.

I acknowledge earnest thanks to my supervisor Ingelin Steinsland for her guidance, support and thoughtful insight. Furthermore I want to thank the student advisor Kristine Lysnes who was always there for any question.

I greatly appreciate and acknowledge the support that I received from every woman helped in distributing and answering my questionnaire. I'm thankful for the cooperation that I received from few organizations and institutions.

I express my gratitude to my beloved husband for his sustainable love, encouragement and cheering that motivated me to keep up the hard work. I would love to thank my beloved son, Josef, for the love, hope and happiness he gives me once I look to his eyes. I am forever indebted to my parents for giving me love, support, opportunities and experiences that have made me who I am. To my friends, thank you for listening, supporting and advising.

Finally I would like to thank myself for keep working hard despite all the challenges were around me. I'm so proud of what I have achieved and learnt during writing this thesis.

Amany El Gharib

November 2019

# Table of Contents

SUMMARY .....	III
PREFACE.....	VI
LIST OF TABLES.....	X
LIST OF FIGURE .....	XII
INTRODUCTION .....	1
WOMEN’S POLITICAL PARTICIPATION .....	6
2.1 POLITICAL PARTICIPATION: .....	6
2.2 WOMEN POLITICAL PARTICIPATION WORLD WIDE .....	8
2.2.1 Historical overview .....	8
2.2.2 Current situation.....	12
2.3 WHY DO WE NEED WOMEN IN POLITICS.....	13
2.4 WOMEN POLITICAL PARTICIPATION IN NORWAY .....	14
2.5 IMMIGRANT WOMEN’S POLITICAL PARTICIPATION IN NORWAY .....	16
SURVEY METHODOLOGY AND QUESTIONNAIRE STRATEGY .....	19
3.1 OVERVIEW OF SURVEY RESEARCH .....	19
3.1.2 Survey Research Methods.....	20
3.1.3 The research planning important steps: .....	21
3.2 DEVELOPING A QUESTIONNAIRE FOR A SURVEY RESEARCH .....	22
3.2.1 Selecting the questionnaire as a research method.....	22
3.2.2 Questionnaire building.....	23
3.2.3 Questionnaire designing.....	24
3.2.4 Questionnaire Formulation: .....	26
3.3 SAMPLING.....	27
3.4 PILOTING .....	28
3.5 QUESTIONNAIRE ANALYSIS .....	29
POLITICAL PARTICIPATION OF IMMIGRANT WOMEN IN NORWAY DURING THE LOCAL ELECTIONS IN 2015 AND THE NATIONAL ELECTIONS 2017 QUESTIONNAIRE .....	30
4.1 THE CONSTRUCTED QUESTIONNAIRE .....	30
4.1.1 The research goals.....	30
4.1.2 The selected research method .....	31
4.1.3 The questionnaire aim.....	32
4.1.4 Targeted group of the questionnaire .....	32
4.1.5 Literature and previous work reviewing .....	32
4.1.6 Questionnaire building.....	33
4.1.7 The questionnaire hypothesis identification: .....	35
4.2 QUESTIONNAIRE ESTABLISHMENT .....	35
4.2.1 Questionnaire design.....	35
4.2.3 Questionnaire administration and spreading.....	40
4.2.4 Sampling .....	41
DATA AND EXPLANATORY ANALYSIS .....	42
5.1 DATA PREPARATION .....	42

5.2	THE COLLECTED DATA DESCRIPTION .....	42
5.3	VARIABLES PREPARATION.....	44
5.4	PERSONAL INFORMATION DATA SET (X1) .....	44
5.4.1	Personal information variables general description .....	45
5.4.2	Correlation .....	51
5.4.3	Collinearity and multicollinearity .....	52
5.4.4	Correlation relation between personal information variables and voting in 2017 national elections variable.....	52
5.4.5	Correlation relation between personal information variables and voting in 2015 local elections variables.....	53
5.5.1	Respondent’s residency variables general description .....	53
5.5.2	Correlation .....	59
5.5.3	Collinearity and multicollinearity .....	59
5.5.4	Correlation relation between residency variables and voting in 2017 national elections variables .....	59
5.5.5	Correlation relation between residency information variables and voting in 2015 local elections variables.....	59
5.6	THE POLITICAL PARTICIPATION DATASET VARIABLES (X3) .....	60
5.6.2	Correlation .....	70
5.6.3	Collinearity and multicollinearity .....	70
5.6.4	Correlation relation between political participation variables and “voting in 2017 national elections” variable.....	70
5.6.5	Correlation relation between political participation variables and “voting in 2015 local elections” variable.....	70
5.7	THE SURROUNDED ENVIRONMENT AND THE POLITICAL VIEW (X4) .....	71
5.7.1	The surrounded environment and the political view variables general description.....	71
5.7.2	Correlation .....	75
5.7.3	Collinearity and multilinear .....	75
5.7.4	Correlation relation between X4 variables and “voting in 2017 national elections variable” 75	75
5.7.5	Relation between X4 variables and “voting in 2015 local elections variable” .....	75
5.8	POLITICAL VIEW DATA SET (X5).....	75
5.8.1	Collinearity and multilinear .....	75
	DATA MODELLING .....	76
6.1	LOGISTIC REGRESSION .....	76
6.1.1	linear regression .....	76
6.1.2	Logistic regression (LR) .....	77
6.1.3	Logistic regression in R .....	78
6.2	PRINCIPLE COMPONENT ANALYSIS (PCA).....	79
6.3	PRINCIPLE COMPONENT LOGISTIC REGRESSION (PCLR).....	81
6.4	STEPWISE MODEL SELECTION.....	82
	DATA ANALYSIS RESULTS .....	83
7.1	RESULTS OF FITTING LOGISTIC REGRESSION MODEL .....	84
7.1.1	With the dependent variable: voting 2017 national elections.....	84
7.1.2	With the dependent variable: voting 2015 local elections .....	90
7.2	PRINCIPLE COMPONENT ANALYSIS (PCA) .....	97
7.3	PRINCIPLE COMPONENT LOGISTIC REGRESSION.....	110
7.3.1	With the dependent variable: voting 2017 national elections.....	110
7.3.2	With the dependent variable: voting 2015 local elections .....	115



7.4	LOGISTIC REGRESSION FOR ALL SIGNIFICANT EXPLANATORY VARIABLES .....	119
7.4.1	Finding the correlation between the significant explanatory variables .....	119
7.4.2	Binomial logistic regression with the dependent variable: voting 2017 national elections	120
7.4.3	Binomial logistic regression with the dependent variable: voting 2015 local elections.....	121
7.5	PRINCIPLE COMPONENT LOGISTIC REGRESSION ALL SIGNIFICANT PRINCIPLE COMPONENTS..	122
7.5.1	Finding the correlation between the significant principle components .....	123
7.5.2	With the dependent variable: voting 2017 national elections.....	124
7.5.3	With the dependent variable: voting 2015 local elections .....	126
	CONCLUSION AND DISCUSSION .....	128
	BIBLIOGRAPHY.....	132
	APPENDIX A.....	141
	APPENDIX B.....	152

# List of tables

7.1	The logistic regression between voting 2017 and x1 data set variables	85
7.2	The predicted probabilities for voting 2017 at each education level	85
7.3	The logistic regression between voting 2017 and x2 data set variables	86
7.4	The logistic regression between voting 2017 and x3 data set variables	87
7.5	The logistic regression between voting 2017 and x5 data set variables	88
7.6	The logistic regression between voting 2015 and x1 data set variables	90
7.7	The logistic regression between voting 2015 and x2 data set variables	91
7.8	The logistic regression between voting 2015 and x3 data set variables	92
7.9	The logistic regression between voting 2015 and x4 data set variables	94
7.10	The logistic regression between voting 2015 and x5 data set variables	95
7.11	The principle components characteristics of x1 data set	97
7.12	The principle components eigenvalues of x1 data set	98
7.13	The principle components' coefficients for x1 data set	99
7.14	The principle components characteristics of x2 data set	100
7.15	The principle components eigenvalues of x2 data set	101
7.16	The principle components' coefficients of the data set x2	102
7.17	The principle components characteristics of x3 data set.	103
7.18	The principle components eigenvalues of x3 data set	104
7.19	The principle components' coefficients of data set's x3	104
7.20	The principle components characteristics of x4 data set	105
7.21	The principle components eigenvalues of x4 data set	106

7.22	The principle components' coefficients for the data set x4	106
7.23	Some of the principle components characteristics of x5 data set	107
7.24	The eigenvalues for some of the principle components for x5 data set	108
7.25	The principle components' coefficients for the data set x5	109
7.26	The logistic regression between voting 2017 and the x1 data set's selected components	111
7.27	The logistic regression between voting 2017 and the x2 data set's selected components	111
7.28	The logistic regression between voting 2017 and the x3 data set's selected components	112
7.29	The logistic regression between voting 2017 and x5 data set's principle components	113
7.30	The logistic regression between voting 2015 and the x1 data set's selected components	115
7.31	The logistic regression between voting 2015 and the x2 data set's selected components	116
7.32	The logistic regression between voting 2015 and the x3 data set's selected components	116
7.33	The logistic regression between voting 2015 and the x4 data set's selected components	117
7.34	The logistic regression between voting 2015 and the x5 data set's selected components	118
7.35	Fitting one logistic regression to all significant variables with voting 2017	121
7.36	Fitting one logistic regression to all significant variables with voting 2015	122
7.37	The logistic regression between voting 2017 and all significant principle components	124
7.38	The logistic regression between voting 2015 and all significant principle components	126
B.1	X5 data set's variables	151

# List of figure

2.1	Women in politics globally in 2019,The source UN women	12
2.2	Norwegian representatives at the country council distributed by gender	14
2.3	Norwegian representatives at local councils distributed by gender	15
2.4	Electoral turnout in Norway segregated by gender	16
2.5	Voting at Norwegian parliament elections distributed by sex and immigration background	17
4.1	The front page of the established questionnaire	36
5.1	The collected data 54 variables”	43
5.2	The “Age” variable distribution among the participants	45
5.3	The ”Social status” variable distribution among the participants	46
5.4	The “Educational status” variable distribution among the participants	47
5.5	The “Employment status” variable distribution among the participants	48
5.6	The “main income of the family” variable distribution among the participants	49
5.7	The “family members number” variable distribution among participants	50
5.8	The “division of household chores” variable distribution among participants	51
5.9	The “ participant’s home country” variable distribution among participants	54
5.10	The “participant’s status in Norway” variable distribution among participants	54
5.11	The “Residency period” variable distribution among participants	55
5.12	The ”Residency permit type in Norway” variable distribution among participants	56
5.13	The “Residency address in Norway” variable distribution among participants	57
5.14	The “Norwegian language level” variable distribution among participants	58
5.15	The “Political participation” variable distribution among participants	60

5.16	The “Why did the respondent vote?” variable distribution among participants	61
5.17	The “Why didn’t the participant vote?” variable distribution among participants	63
5.18	The “Voting at the election in home” variable distribution among participants	64
5.19	“A member or a candidate of a political party” variable distribution	64
5.20	“A member of an interest group working for social or political change” variable distribution among participants	65
5.21	“A volunteer for a group or organization other than a political party” variable distribution among participants	66
5.22	“Taking a part in a demonstration, in the last 12 months” variable distribution	67
5.23	“Signing a petition, in the last 12 months” variable distribution among participants	68
5.24	“Voting in the 2015 local elections” variable distribution among participants	69
5.25	“Voting in the 2017 national elections” variable distribution among participants	69
5.26	“How often did your family used to discuss politics” variable distribution among participants	71
5.27	“How often do you discuss politics now?” variable distribution among participants	72
5.28	“How many of your friends have an immigrant background? “ variable distribution	73
5.29	“Pressure from your with your choice of a political party” variable distribution	74
7.1	Scree plot of x1 data set’s principle components	98
7.2	The biplot of pc1, pc2 and the explanatory variables of x1 data set	100
7.3	Scree plot of x2 data set’s principle components	101
7.4	The biplot of pc1, pc2 and the explanatory variables of x2 data set	102
7.5	Scree plot of x3 data set’s principle components	103
7.6	The biplot of pc1, pc2 and the explanatory variables of x3 data set	105

7.7	Scree plot of x4 data set's principle components	106
7.8	The biplot of pc1, pc2 and the explanatory variables of x4 data set	107
7.9	Scree plot of data set's x5 principle components	108
7.10	The biplot of pc1, pc2 and the explanatory variables of x5 data set	110

# Chapter 1

## Introduction

### Populations Statistics in Norway

Statistics Norway (*statistisk sentralbyrå (SSB)*)<sup>1</sup> defines the immigration population as *people who haven't neither parents nor grandparents born in Norway*. Based on this definition *the immigrants' group is divided to 2 generations; first generation who moved to Norway without Norwegian background, and second generation who was born in Norway, but by foreign-born parents*.

Based on *statistisk sentralbyrå (SSB)* <sup>2</sup>, the percent of immigrants in Norway is 14.4% of the total population at the beginning of the year 2019 while the percent of the Norwegian-born to immigrant parents represents 3.4 % of the total population. In numbers, there are 765 100 immigrants and 179 300 Norwegian-born to immigrant parents live in Norway. Children with parents' Pakistani background represent the highest number, with 17 300. While the Norwegian-born to immigrant parents with Somali background represent the second largest group, with 14 200. The growth of immigrants' number for the year 2018 is 18 400 which make it the lowest since 2005.<sup>3</sup>

In more general perspective, the number of the total population is approximately 5.3 million in the year 2019, divided to 2,69 million men and 2.64 million women living in Norway, based on *The Statistics Portal* <sup>4</sup>

---

<sup>1</sup> Byberg, Ingvild. 2002. Immigrant Women in Norway A Summary of Findings on Demography, Education, Labour and Income.

<sup>2</sup> Statistics Norway. 2019. Lower Growth in the Number of Immigrants.

<sup>3</sup> Steinkellner, Alice, Statistics Norway- Statistisk sentralbyrå., SSB 2019. The Growth in the Number of Immigrants in Norway Is the Lowest for a Number of Years. The Increase Was 18 400 in 2018, Making This the Smallest Growth since 2005.

<sup>4</sup> Statista. 2019. *Population in Norway from 2009 to 2019, by Gender*.

But what do we know about immigrant women in Norway? Let us first wonder if there is segregated data basically represents the immigrant women at the population statistics or diagrams! The populations statistics in Norway is classified based on gender “women and men “, age categories, or even immigration principle. But there is little detailed statistics represent immigrant women in Norway. The gender segregated data and statistics is built based on gender “ women or men”. On the other hand, the immigrants” segregated data is built based on categorized mainly based on immigrant and non-immigrant background.

There are a very few publications <sup>5</sup> about gender equality at the Norwegian society with detailed statistics and information about immigrant women in Norway <sup>6</sup>. These aren't updated regularly and even don't cover all the related issues and society aspects such like politics, cultural issues , social issues ,....etc. At the end, what we do know about immigrant women in Norway is so little<sup>7</sup>. There is no much updated detailed statistics represent them or related to their issues and needs. And even we don't know clearly and deeply what is behind the numbers at these statistics and figures.

This thesis aims to gain understanding of the political participation for immigrant women in Norway during the local elections 2015 and the National election 2017. Modelling and inference is done using the logistic regression and principle component regression to highlight the main factors which affect this political participation.

The approach t used is divided to seven main steps which are (1) building a statistical data collection method which is a questionnaire, that is constructed based on previous similar studies” will be discussed in details in chapter 3 and 4” . (2) analysis of the collected data using the binomial multiple logistic regression. (3) dimensions reduction for the data explanatory variables using principle components analysis (4) Data analysis using the principle components logistic regression (5) Gathering the significant explanatory variables resulted from the binomial logistic regression to fit a new binomial logistic regression. (6)

---

<sup>5</sup> Høstmark, Maria-SSB. 2013. Velgerundersøkelse Blant Personer Med Innvandrerbakgrunn 2013.

<sup>6</sup> Dalgard, Anne Berit. 2018. Levekår Blant Norskfødte Med Innvandrerforeldre i Norge 2016.

<sup>7</sup> Karin Hamre (ed.) et al. 2018. Women and Men in Norway. [https://www.ssb.no/en/befolkning/artikler-ogpublikasjoner/\\_attachment/347081?\\_ts=1632b8bcb0](https://www.ssb.no/en/befolkning/artikler-ogpublikasjoner/_attachment/347081?_ts=1632b8bcb0).



gathering the significant principle components resulted from the principle components logistic regression to fit a new principle components logistic regression. (7) finding the relation between the significant explanatory variables resulted in step (2) and the significant principle components resulted in step (4).

The main challenges here are building the questionnaire and distributing it. Since it was distributed electronically, the challenge of its distribution that the sample is biased already even before starting the analysis. The second main problem that large number of immigrant women can't communicate very well using the Norwegian language so it was written in both Norwegian and English but still there are a lot of them can neither read Norwegian nor English.

In this thesis a statistical quantitative data collection method is built based on previous existence questionnaires and studies at the same field. The research methodology, questionnaire designing and building are previewed and applied during designing and building the questionnaire structure. This questionnaire is released online. And the collected data were used at the data analysis step.

The case study of the statistical study here targets the immigrant women in Norway to learn more about their political participation in Norway by using the measure tool “ voting in 2017 national elections” and “ voting in 2015 local elections”. The aim of this case study is to address the factors that affect their political participation and the obstacles that challenge their political activities.

Chapter 2 discusses women's political participation. It defines political participation, and the factors that influences people's political participation. A historical overview of the struggle of women's rights movement is presented. Why it is important having women in politics and the current situation of women's political participation in Norway is reviewed. Finally, some background about immigrant women's political participation in Norway is presented

Chapter 3 discusses the research methodology and the questionnaire strategy. The research methodology, the survey research and classifies the research based on its purpose are presented as well. Presenting the research different methods, the essential steps of the research planning, the process of developing the questionnaire for the survey research. The selecting of the questionnaire as a research method, the questionnaire building steps and the

questionnaire design phase is summarized here. The questionnaire formulation, the questionnaire formatting phases, the questionnaire the sampling, piloting and analysis is discussed in addition.

Chapter 4 presents practical implementation to the theories in chapter 3. This chapter shows the steps of designing , constructing, formatting, formulating and spreading our questionnaire. It shows as well how and why it the questionnaire is built based on previous studies.

Chapter 5 shows explanatory analysis to the data explanatory variables. Since the built questionnaire has five main sections. In data analysis, each section is considered as a data set. The explanatory analysis is applied to each data set aside. Exploring the variables plus to check the correlation between the variables within the data set, multicollinearity and correlation with the dependent variables are discussed here as well.

Chapter 6 presents theoretical background about the logistic regression, principle components analysis and principle components logistic regression. Standardising data, data imputation and model selection technique are discussed as well.

Chapter 7 is the core chapter where all the analysis is presented there. The data analysis phases are divided to 5 phases; taking into account that the data contains 52 explanatory variables and 2 dependent variables. And the number of the completed observations is 95.

The first phase is achieved by analysing the data explanatory variables using the binomial logistic regression statistical method. The binomial logistic regression is used here since we have 2 binary dependent variables and this is the ideal method could be applied at this case. It is applied to each data set aside, furthermore the applied model selection technique is “ the backward forward” method.

The second phase is implemented a data dimensional reduction technique using principle components analysis. This technique is implement to each data set aside also.

The third phase is used the selected principle components for each data set to fit a principle components binomial logistic regression model using the “backward forward” model selection technique. Fitting this model is applied to each data set aside for both responses, each one again aside. The fourth phase is achieved by using the significant explanatory variables that will be resulted from fitting the binomial logistic regression in the first phase (section 7.1) to each data aside. In this phase the significant variables is used to fit one new

binomial logistic regression model. Comparing the results with the results of the first phase as well is implemented.

The fifth phase is used the significant principle components that is resulted from fitting the principle components binomial logistic regression in the third phase ( section 7.3) to each data set a side, to fit one new principle components binomial logistic regression model and compare the results by the results of third phase.

Chapter 8 summarizes and discusses the results.

The constructed questionnaire is presented in appendix A.

# Chapter 2

## Women's political participation

### Summary

This chapter discusses women's political participation. At the first section the political participation is defined, and the factors that influence people's political participation are introduced. A historical overview of the struggle of women's rights movement is presented in the first part of the second section. Highlighting the current situation of women's rights and political participation at the world is mentioned at the second part of the second section. Why it is important having women in politics is overviewed at the third section. While the current situation of women's political participation in Norway is mentioned at the fourth section. Finally, some background about immigrant women's political participation in Norway is presented at the fourth section.

### 2.1 Political participation:

Defining the political participation as the set of activities that are implemented by the government, politicians, diplomatic actors and the political organs, is considered as a limited definition.<sup>8</sup>

From a wider perspective the political participation is defined as the range of activities that are achieved at different levels. Where the ordinary citizens, people and individuals contribute to. Their contribution goal is to develop and express their opinions, beliefs and thoughts to influence the decisions that shape the policies. These policies are taken by powerful actors; politicians, governmental organs, political parties, organizations or even groups" at the society and affect their lives.

---

<sup>8</sup> Khasnabis C, Heinicke M, Motsch K, Achu K. 2010. *Community-Based Rehabilitation: CBR Guidelines*. Geneva: World Health Organization;

Voting has been since along time the primary and powerful tool for letting individuals expressing their opinions and make their voices heard. Furthermore, voting turnout was considered as a measure tool for citizens' political participation in some countries such like the United states. <sup>9</sup>

The political participation topology that was suggested by Teorell et al. (2007) <sup>10</sup> consists of five dimensions which are; electoral participation, consumer participation, party activity, protests activity and contact activity

The first dimension presents going to the ballots procedure. While if the citizen participate in signing a petition, donating money to charity or boycotting, this means that he contributes in the political consumption. When the individual is a member , volunteer, employee, activist within a political party or even donating money to a political party, his participation classified under the party activity dimension. Protest activity covers taking a part in a demonstration, protest or a strike. Contact activity includes contacting politicians, civilian servant or general organizations.

The individuals in the society can participate politically to influence the public polices, decisions and taking a part of the political representation locally and internationally. But there are always obstacles, challenges and barriers in front of ordinary people in their political participation. There barriers can be; <sup>11</sup>

- ❖ Financial situation: people who has limited income, are poor or even don't have any income, mostly focus on fixing their needs before the participation in politics. Survival activities come first at their priorities list. This leads to the limitation in their time, effort or interest to participate in politics

---

<sup>9</sup> Ekman, Joakim, and Erik Amnå. 2012. "Political Participation and Civic Engagement: Towards a New Typology." Human Affairs.

<sup>10</sup> Ekman, Joakim, and Erik Amnå. 2012. "Political Participation and Civic Engagement: Towards a New Typology." Human Affairs.

<sup>11</sup> Khasnabis C, Heinicke Motsch K, Achu K. 2010. Community-Based Rehabilitation: CBR Guidelines. Geneva: World Health Organization;

- ❖ Education: is a primary obstacle for having meaningful political participation. People need knowledge and information about the daily actions happen around them. So they can be able to create their own political attitudes and opinions.
- ❖ Personal reasons: people don't believe in politics, don't have motivation, don't trust politicians, or even don't trust that they could contribute in doing a public change. From the other hand some people have personal responsibilities and priorities that take the majority of their efforts and time. So they aren't interested in political participation.
- ❖ Social isolation: that results in a limited network to the isolated people, that leads to obstruct their motivation and participation.
- ❖ Discrimination and stereotypes: based on sex, language, colour, religion or ethnicity affect the minorities political participation.
- ❖ Legal obstacles: people under 18 years old can't vote for example . People who are not citizens and live in the host countries since a while can't vote as well.
- ❖ Lack of role models: people with disabilities are encourage by role models with disabilities . Immigrant especially immigrant women are motivated by immigrant female role models .

## **2.2 Women political participation world wide**

### **2.2.1 Historical overview**

Despite that the duties, responsibilities and rights should be shared equally by all citizens at the society, this hadn't happened at the past. It is said that historically, women and minorities at the majority of the western countries were excluded from the citizenship right. They weren't considered as citizens but were considered as property either as slaves or domestic property. This deny had continued for thousands of years until women started to argue and demand their rights loudly at the beginning of the nineteenth century. Even when these groups have been granted the citizenship right, they were second class citizens which means

that the oppression and inequality they had suffered didn't stop completely as they haven't been granted all of their rights<sup>12</sup>.

It is said that women's political and liberation movement for gaining their rights had started at the third part of the eighteenth century, but the actual foundation of this movement was at the beginning part of the nineteenth century<sup>13</sup>.

At that time and before, all doors at the society were closed in front of women. Women had to be subordinators to men not only at real life. But also, men used to aggressively attack women in their publishes, and so the political movement for women's rights was a response to raise women's voices not only against these published works by men, but moreover to raise the discussions about their rights and citizenship. Simply what we call today "gender". Feminists (women's rights activists) at that time confirmed that it's not the nature which had made women's inferior, but it is the culture at the societies.<sup>14</sup>

Women had participated actively at the subsistence riots during the French revolution. Collective voices started to demand women's access to education, the elimination of discriminatory laws and even women's representation in the Estate general. These demands were collected at the pre-revolutionary petitions. Nevertheless, the national assembly excluded women's political representation during the revolution by denying their access to political sovereignty. In 1791 a declaration of the Rights of Women was published by Olympe de Gouges, who demanded women's legal existence as citizens at the sovereign nation, but at the same time called for the exclusion of women from the political representation. <sup>15</sup>

---

<sup>12</sup> Lister, Michael, and Emily Pia. 2008. Citizenship in Contemporary Europe *Citizenship in Contemporary Europe*.

<sup>13</sup> Website, Scholastic. "The History of Women's Suffrage." <http://teacher.scholastic.com/activities/suffrage/history.htm>.

<sup>14</sup> Debbie Wigglesworth, Jalna Hanmer, Catherine Euler, Val Balding, Teresa Ortíz Gómez, Candida Martínez López, Margarita M. Birriel Salcedo, Pilar Ballarín Domingo, Hildur Ve, Berit Bareksten, Nicky Le Feuvre, Eeva Raevaara, Susanna Taskinen. "Women in the European Union." 2015. <http://www.helsinki.fi/science/xantippa/wee/wee24.html>.

<sup>15</sup> Debbie Wigglesworth, Jalna Hanmer, Catherine Euler, Val Balding, Teresa Ortíz Gómez, Candida Martínez López, Margarita M. Birriel Salcedo, Pilar Ballarín Domingo, Hildur Ve, Berit Bareksten, Nicky Le Feuvre, Eeva Raevaara, Susanna Taskinen. "Women in the European Union." 2015. <http://www.helsinki.fi/science/xantippa/wee/wee24.html>.

In the English-speaking world, the struggle against the unequal rights and discrimination against women is thought that it had begun with the publications of Mary Wollstonecraft who focused in her discussion on that the imbalance between sexes was not due to biological differences, but due to women's lives, socialization and education. Her argument discussed the constrictions on women's lives such as the social orders that were made by men to stop women freely express their abilities. Mary Wollstonecraft's publication *Vindication of the Rights of Woman* (1793) was a key and a starting point for the later women's liberation movements. She saw the education as an appropriate mean to balance the inequality between women and men and to achieve women's autonomy. Her equality plane and the argument she focused on with some small formulaic nuances was a starting point to the feminist movement in Europe in the nineteenth and even twentieth centuries.

Women's rights activists focused on women's right to vote as a pivotal right. Female's voting was a principle tool which would let them enter to the decision-making positions to change the laws and legislate new laws that could guarantee abolishing the social inequality. This was the main cause of feminist mobilization in the late nineteenth and early twentieth century.<sup>16</sup>

Women's suffrage demands were resisted aggressively by authorities and societies, and the radical mentality and cultural rules of that period is a good explanation for that. The suffrage right was considered as a threat to family, home and the whole society. Despite all of that women became increasingly active in their quest to their own suffrage right at the end of the ninth century, but the road to suffrage and ballots was not easy, of course it was full of big pitfalls and small victories before the suffrage was finally gained.<sup>17</sup>

Women gained the suffrage on the national level in New Zealand in 1893. It was followed by Australia in 1902. Women were granted the voting right early in the 20th century in the European countries such as Finland followed in 1906, Norway in 1913 and Denmark and

---

<sup>16</sup> Debbie Wigglesworth , Jalna Hanmer , Catherine Euler, Val Balding, Teresa Ortíz Gómez, Candida Martínez López, Margarita M. Birriel Salcedo, Pilar Ballarín Domingo, Hildur Ve, Berit Bareksten, Nicky Le Feuvre, Eeva Raevaara, Susanna Taskinen. "Women in the European Union." 2015. <http://www.helsinki.fi/science/xantippa/wee/wee24.html>.

<sup>17</sup> Debbie Wigglesworth , Jalna Hanmer , Catherine Euler, Val Balding, Teresa Ortíz Gómez, Candida Martínez López, Margarita M. Birriel Salcedo, Pilar Ballarín Domingo, Hildur Ve, Berit Bareksten, Nicky Le Feuvre, Eeva Raevaara, Susanna Taskinen. "Women in the European Union." 2015. <http://www.helsinki.fi/science/xantippa/wee/wee24.html>.



Iceland in 1915. Women at the other continental powers, The United States, Great Britain and Canada, didn't achieve the same rights until the end of World War I.<sup>18</sup>

In 1917 women in the Union of Soviet Socialist Republics and Netherlands accorded suffrage in 1917. On the other hand Austria, Czechoslovakia, Poland, and Sweden the suffrage was granted in 1918; followed by Germany and Luxembourg in 1919. Women couldn't achieve ballots in Great Britain until 1928. The Spanish women gained the suffrage in 1931 while the French women waited until 1944. Belgium, Italy, Romania, and Yugoslavia accorded it in 1946. Switzerland granted it in 1971 while women in Liechtenstein waited until 1984.

In 1916, the Canadian women gained the suffrage in Alberta, Manitoba, and Saskatchewan, women at other provinces won the suffrage right after the federal in 1918, the last province to grant women suffrage was Quebec and that in 1940.

Women were active since the period of nineteenth and early twentieth century to win the suffrage. After the end of the World War I and World War II, women gained the suffrage in different countries in Latin America, Asia, Africa and Middle East North Africa region.<sup>19</sup>

Finally, and as a conclusion, women won the suffrage in the northern countries such like Norway and Finland as the first European countries who committed women equal political rights and granted them suffrage. The Norwegian women's movement which was born in 1893 was a key reason led to political equality at early time in comparison with other countries.

Women's political struggle in shaping the policies and form the public decisions is not limited only in the suffrage, women started to run after the political representation and participation in parliaments, government, political parties, state organizations nationally and internationally.<sup>20</sup>

---

<sup>18</sup> Website, Scholastic. "The History of Women's Suffrage." <http://teacher.scholastic.com/activities/suffrage/history.htm>.

<sup>19</sup> Website, Scholastic. "The History of Women's Suffrage." <http://teacher.scholastic.com/activities/suffrage/history.htm>.

<sup>20</sup> Debbie Wigglesworth, Jalna Hanmer, Catherine Euler, Val Balding, Teresa Ortíz Gómez, Candida Martínez López, Margarita M. Birriel Salcedo, Pilar Ballarín Domingo, Hildur Ve, Berit Bareksten, Nicky Le Feuvre, Eeva Raevaara, Susanna Taskinen. "Women in the European Union." 2015. <http://www.helsinki.fi/science/xantippa/wee/wee24.html>.



Women who served as head of states in June 2019 were only 12, while who served as a head of the government were 11 women. The highest number of parliamentarian women's representation worldwide is 61.3% and Rwanda is the country that has this percentage. Globally, women represent less than 10% of the (single or lower houses) at the parliament in 27 countries.

### **2.3 Why do we need women in politics**

According to The National Democratic Institute for international affairs, Women's empowerment in politics will result in decreasing the gender gap in the political positions and enable them to participate in policies shaping and decisions making which will be reflected as a progress and improvement of different life aspects.<sup>22</sup>

The seventieth Secretary-General of the United Nations Kofi Annan noted, "study after study has taught us, there is no tool for development more effective than the empowerment of women. No other policy is as likely to raise economic productivity or to reduce child and maternal mortality. No other policy is as sure to improve nutrition and promote health, including the prevention of HIV/AIDS. No other policy is as powerful in increasing the chances of education for the next generation."

"But whatever the very real benefits of investing in women, the most important fact remains: women themselves have the rights to live in dignity, in freedom from want and freedom from fear", he continued. That confirms the undeniable role for women at the society and the emergent need for empower them. Some of these studies; women's Assessments of Gender Equality<sup>23</sup>, Gender, gender roles, and gender identity<sup>24</sup>, Gender Equality – Why It Matters."UN women"<sup>25</sup>

---

<sup>22</sup> The National Democratic Institute for international affairs. "WHY WOMEN IN POLITICS?"

<sup>23</sup> Kurzman, Charles et al. 2019. "Women's Assessments of Gender Equality." Sociological Research for a Dynamic World.

<sup>24</sup> Roles, Gender, and Gender Identity. 1999. "Gender, Gender Roles, and Gender Identity." Identity.

<sup>25</sup> Women, UN. 2016. "Gender Equality – Why It Matters." Sustainable Development Goals.

## 2.4 Women political participation in Norway

In this section we present background statistics for the members of Norwegian parliament and local councils in Norway for men and females. Furthermore the participation in elections, i.e. election turn out. The background data finds that <sup>26</sup>;few women were sat in the Norwegian parliament and local councils at the period between the period after world war II and the sixties (1945-1960). The numbers have started to raise at late of the sixties significantly and the percentage has inflated from less than 10% in 1969.

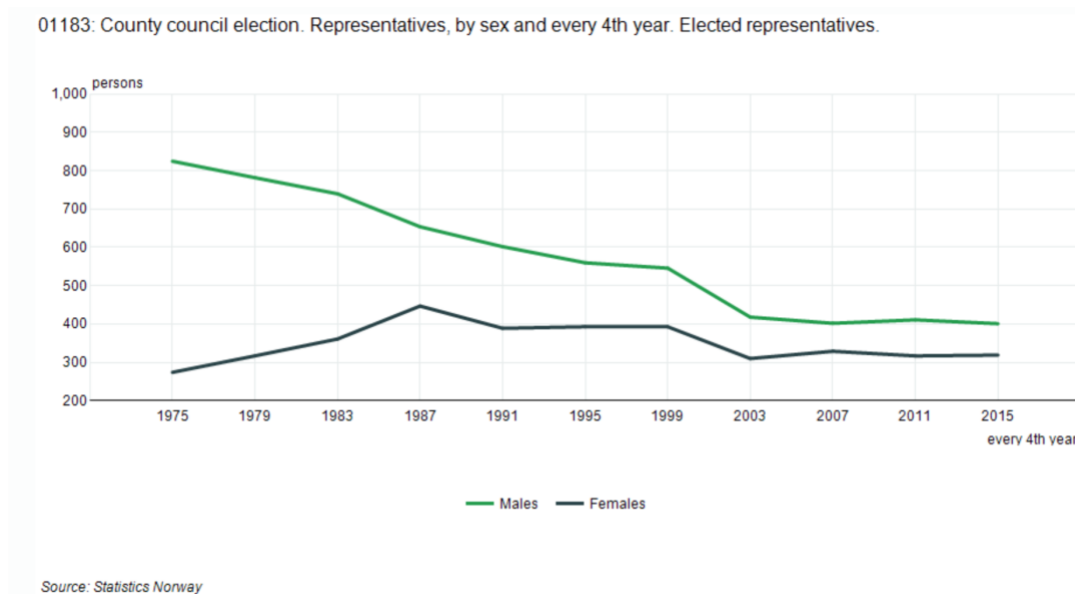


Figure ( 2.2): Norwegian representatives at the country council distributed by gender, the source: [www.ssb.no](http://www.ssb.no)

For the female representatives' number at the country council, contained raising to around 40% in 1991 with the highest peak in 1987 which reached around 45 % The percentages remained stable until 1999 and then went down to around 30 % since then to the last elections in 2015 stayed in a line swing around 33% and 30%, figurer (2.2).

The female representatives' number at the municipal council contained as well in growing to around 32% in 1983 with the highest peak in 1987 which reached around 43 % The

<sup>26</sup> Karin Hamre (ed.) et al. 2018. Women and Men in Norway. [https://www.ssb.no/en/befolkning/artikler-og-publikasjoner/\\_attachment/347081?\\_ts=1632b8bcba0](https://www.ssb.no/en/befolkning/artikler-og-publikasjoner/_attachment/347081?_ts=1632b8bcba0).

percentages went down again in 1991 to 38% remained and then again to 42 % in 1995. The percentage since then until the last elections in 2015 has fluctuated around 40% (figure 2.3).

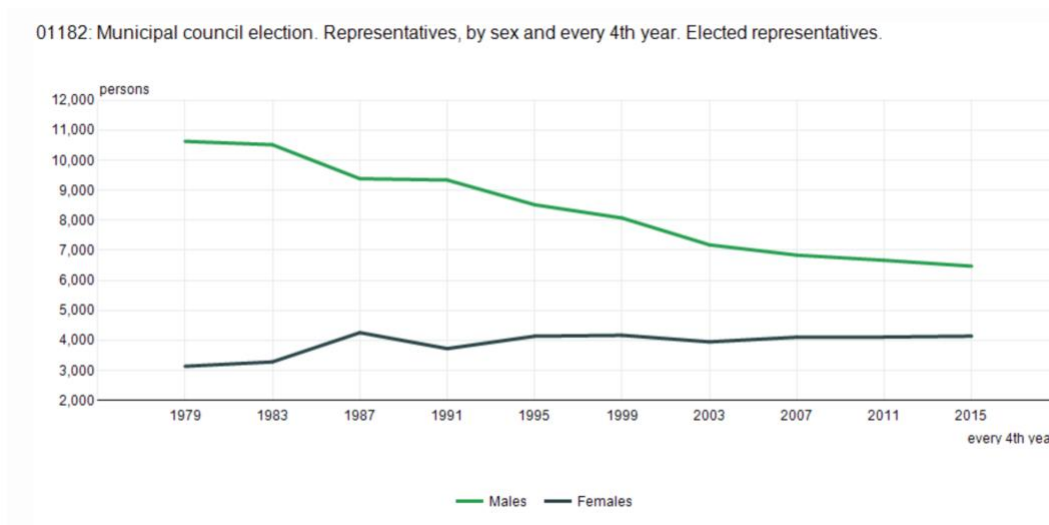
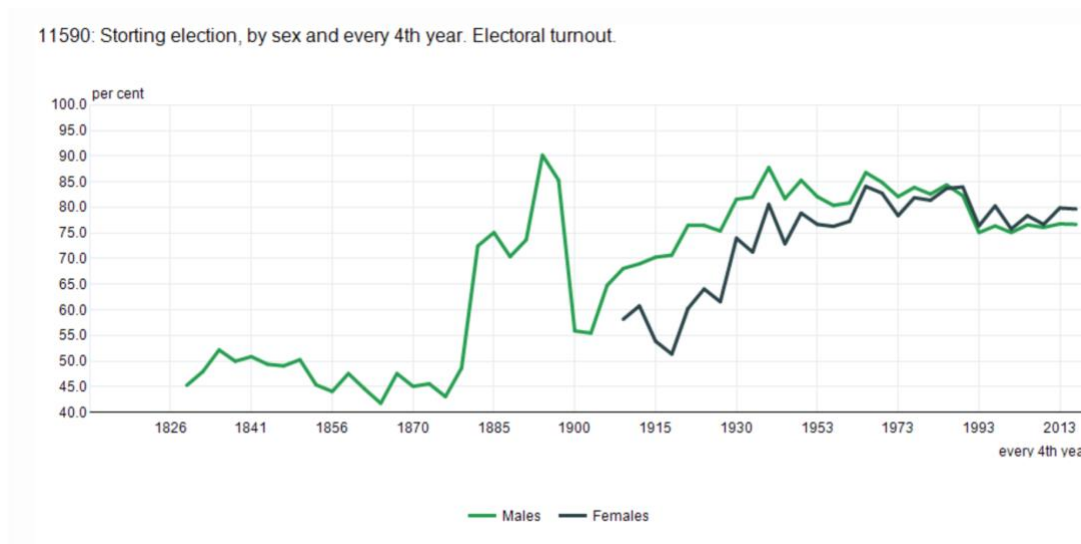


Figure (2.3): Norwegian representatives at local councils distributed by gender, the source: [www.ssb.no](http://www.ssb.no)

An interesting point here is that women's turnout has changed historically. Based on Statistisk sentralbyrå - Statistics Norway (SSB) at the first elections, directly after the world war II the percentage of men was higher than women a little bit. The imbalance has been rectified by 1980s and more women participated at the voting than men. 80% of women used their right of voting in the last elections 2017 while men's turnout rate was 77% of men in Norway<sup>27</sup>.

<sup>27</sup> Karin Hamre (ed.) et al. 2018. Women and Men in Norway. [https://www.ssb.no/en/befolkning/artikler-og-publikasjoner/\\_attachment/347081?\\_ts=1632b8bcba0](https://www.ssb.no/en/befolkning/artikler-og-publikasjoner/_attachment/347081?_ts=1632b8bcba0).



Figure( 2.4): Electoral turnout in Norway segregated by gender, the source: www.ssb.no

## 2.5 Immigrant women’s political participation in Norway

The political participation for immigrants in Norway recorded low rates since 1997, which is the first year when Statistics in Norway started to record and measure immigrants’ political participation<sup>28</sup>.

At the Norwegian parliament elections in 2013, 103 000 men and 112000 women with immigrant background were eligible to participated at the elections. At these elections 78% of the whole population voted, while among Norwegian with immigrant background 53% who voted. There are only small differences between the last four elections based on SSB statistics (Figure 2.5)

2013 is counted as a special year for women’s political voting at the elections as the numbers of women, with or without immigrant background, who used their voting rights have exceeded the number of men who voted. (Figure 2.5)

<sup>28</sup> Hoveid, Kristin Sandnes, Toril. 2015. Innvandrere Og Norskfødte Med Innvandrerforeldre i et Kjønn- Og Likestillingsperspektiv Utdanning, Arbeid Og Samfunnsdeltakelse.

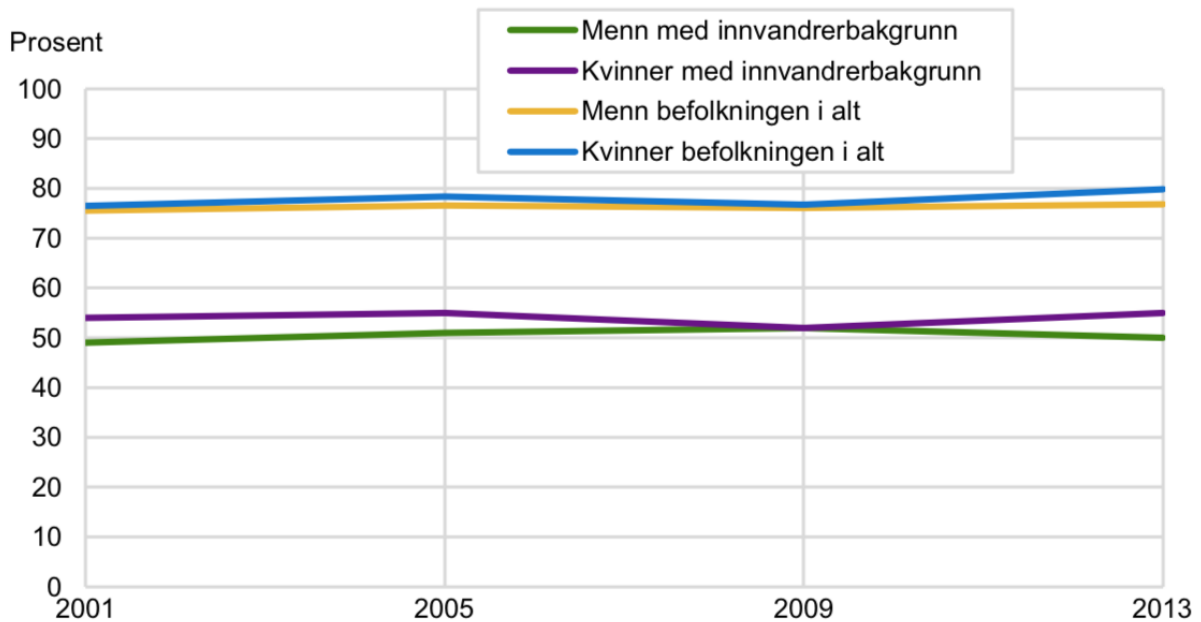


Figure (2.5): Voting at Norwegian parliament elections distributed by sex and immigration background, the source: [www.ssb.no](http://www.ssb.no) ; through “Innvandrere og norskfødte med innvandrerforeldre” report for 2015

The percentage of Norwegian men with immigrant background who voted at the parliament elections declined from 52 % in the 2009 elections to 50% in 2013 elections. Nevertheless, immigrant women’s percentage raised from ? to ? . Furthermore the difference between women and men in the voting turnout is greater with immigrants background than without immigrants background . But still almost half of immigrants don’t use their right to vote.

On the other hand, at the municipal councils’ election in 2011 was 196000 men and 191000 women with immigrant background eligible to participate at the elections. Women voting turnout was greater than men generally.

Moreover the immigrant women’s voting turnout is greater than men at these elections.. The percent of Norwegian women citizens who voted was 44% while the percent of Norwegian men was 41%. Furthermore, the percent of women with foreign citizenship is 35 % while among men with foreign citizenship was 29% .<sup>29</sup> ( Norwegian citizenship is not an obligation to vote at this election)

---

<sup>29</sup> Hoveid, Kristin Sandnes, Toril. 2015. Innvandrere Og Norskfødte Med Innvandrerforeldre i et Kjønn- Og Likestillingsperspektiv Utdanning, Arbeid Og Samfunnsdeltakelse.

It is obvious that women's political participation is important moreover the political participation of immigrant woman. But what do we know about immigrant women's political participation is still so little, what is behind these numbers ? what does affect their political participation? What are the challenges in front their participation and what is the motivation behind their participation?

This thesis tries to contribute in answering these questions using statistical methods of binomial multiple logistic regression and principle components logistic regression.



# Chapter 3

## Survey methodology and questionnaire strategy

### Summary

This chapter is divided to two main sections. The first section defines the research methodology. It presents the survey research and reviews the research different methods as well. Discussing the essential steps of the research planning is a part of this section

The second section discusses the questionnaire strategy, by presenting the process of developing the questionnaire for the survey research. It summarises the selecting of the questionnaire as a research method, reviews the questionnaire building steps and the questionnaire design phase. It discusses the questionnaire formulation and the questionnaire formatting phases. And presents the questionnaire the sampling, piloting and analysis

### 3.1 Overview of survey research

#### 3.1.1 What is the survey and the survey research?

One of the approaches to study a phenomenon, problem or an aspect at the society is to ask the population about it. To know more about people's satisfying of public services, their opinions about political parties or even what do they think about unemployment rates at the society, you need to contact people and ask them about this, and this what we call a survey. A survey is defined in different ways. Simply it is a set of questions that are asked to a predefined group of people at the society. Asking them about their opinions, attitudes, believes, assessment, knowledge or even what do they think about a specific subject<sup>30</sup>.

---

<sup>30</sup> Ponto, Julie. 2015. "Understanding and Evaluating Survey Research." Journal of the advanced practitioner in oncology.

Reaching the whole population is a very difficult and unachievable at the practical level. Furthermore, it consumes a lot of time and is very expensive in addition. Alternatively, the survey is one of the best method to data collection. It targets a pre-defined targeted group of people , that is called sample, which is qualified to represent the whole population. This method saves time, efforts and money.

For using the survey results in shaping polices, drawing decisions or summarizing conclusions, it should be built in a scientific way. People's thoughts, knowledges, attitudes and daily practises can be translated scientifically to policies, laws and decisions using the survey research.

### 3.1.2 Survey Research Methods

The research can be achieved using different methods, for example such as: <sup>31</sup>

- **Face to face interviews:** The researcher contacts the people directly face to face at the street, home, shop, university or organisation. It is an effective method as the researcher can sell his research to the respondent so he will have more respondents. This method consumes a lot of time, expensive and in addition impractical when large sample is required. There is a bias margin since the researcher writing the respondent's answers not the respondents himself.
- **Telephone interviews<sup>32</sup>:**  
The researcher here call people on phone for having the interview. It is a two-sided interaction on the phone. It is more flexible method, faster, less expensive and the

---

<sup>31</sup> Kelley, Kate, Belinda Clark, Vivienne Brown, and John Sitzia. 2003. "Good Practice in the Conduct and Reporting of Survey Research." *International Journal for Quality in Health Care*.

Qu, Sandy Q., and John Dumay. 2011. "The Qualitative Research Interview." *Qualitative Research in Accounting and Management*.

<sup>32</sup> Roopa, S, and MS Rani. 2012. "Questionnaire Designing for a Survey." *The Journal of Indian Orthodontic Society*.

researcher can cover wider geographical areas and asking complex questions as well in some cases. Nevertheless, it can record higher respondents but people feel less embarrassed to refuse the participation at the interview

- **Questionnaires<sup>33</sup>:**

It is a series of questions which focuses on the research goals and objectives. The main challenge here is that this method depends on the respondents understating and interpretation to the questions which he will reflect in his answers, on the other side as the respondent answer by himself this eliminate the bias of the interviews method. Complex questions here is a big risk. On the other hand, the participation here is anonym which gives the space to ask more sensitive questions if it is needed and encourage more people to respond as their identities are unknown.

Based on the way this questionnaire can be spread and reached, it could be classified as:

Postal Questionnaire: is printed and sent by mail to the respondents usually without any previous contact between the researcher and the respondent.

Online Questionnaire<sup>34</sup>: This method is an outcome of the rapid growth of technology at the modern time. It is a fast method, with wider geographical reach. Could be low cost at some cases. Efficient with direct entry. The main challenge here is the bias.

### **3.1.3 The research planning important steps<sup>35</sup>:**

- ❖ Determining the research goals, questions and problems. At this step, the researcher should pay attention to have knowledge about what he is looking for

---

<sup>33</sup> Laaksonen, Seppo. 2018. Survey Methodology and Missing Data: Tools and Techniques for Practitioners Survey Methodology and Missing Data: Tools and Techniques for Practitioners.

<sup>34</sup> Ball, Helen L. 2019. "Conducting Online Surveys." *Journal of Huma*

<sup>35</sup> Kelley, Kate, Belinda Clark, Vivienne Brown, and John Sitzia. 2003. "Good Practice in the Conduct and Reporting of Survey Research." *International Journal for Quality in Health Care*.

and searching about. He has to widen his experience ground at this field by reviewing previous studies, researches and literature at this field.

- ❖ Detailed literature preview to some of the previous studies and researches at the subject and related topics to it
- ❖ Naming the targeted group at this research; and collecting information about them, learn their language in case the researcher has to contact them directly and they don't understand his language. How they will be reached
- ❖ Counting of the available resources that the researcher can employ to achieve the research goals (human resources, financial resources...etc.)
- ❖ Put a time plane and milestones for achieving the research goals.
- ❖ Choosing the suitable research method to collect the required data, based on the mentioned steps above.
- ❖ Data collection
- ❖ Data Analysis
- ❖ Reporting and final conclusions and recommendations.

## **3.2 Developing a questionnaire for a survey research**

### **3.2.1 Selecting the questionnaire as a research method**

Before choosing the questionnaire as a research method, the researchers have to ask themselves these questions <sup>36</sup>:

- ❖ Is the questionnaire the best way to collect the information for this research? Is it the most suitable method for the situation? why?
- ❖ Will the questionnaire add to the data analysis? What? How?
- ❖ Will the questionnaire address the research problem and contribute in finding answers to the research question? How?
- ❖ Will the questionnaire be formulated to reach the point at the previous question? How?

---

<sup>36</sup> Colosi, Laura. 2006. Cornell Cooperative Extension Designing an Effective Questionnaire.

- ❖ Is it clear to the researcher what is the targeted respondents to this research methodology?
- ❖ Will researcher be able to reach the respondents? How?
- ❖ Will the respondents have clear and enough information to be able to answer the questionnaire? How?
- ❖ Will the questions be clear enough to the targeted group?
- ❖ Is the targeted group willing to participate and answer?

If all the answers to the previous polar (yes/no) questions are “YES”, and if the researchers find answers to the questions followed the polar (yes/no) questions, there is a strong and an enough evidence that has made the questionnaire the best research method for the survey research <sup>37</sup>.

### **3.2.2 Questionnaire building<sup>38</sup>**

- ❖ Defining the questionnaire hypothesis, assumptions and purpose which harmonise with the conducted study objectives.
- ❖ Determining the targeted group by this questionnaire “who are the respondents”, and how will you reach them?
- ❖ Designing and planning the questionnaire
- ❖ Writing the questionnaire
- ❖ Pilot study
- ❖ Sampling
- ❖ Administrating the questionnaire
- ❖ The analysis of the questionnaire
- ❖ The results understanding and interpreting
- ❖ Reporting and documentation

---

<sup>37</sup> Laaksonen, Seppo. 2018. *Survey Methodology and Missing Data: Tools and Techniques for Practitioners Survey Methodology and Missing Data: Tools and Techniques for Practitioners*

<sup>38</sup> Boynton, Petra M., and Trisha Greenhalgh. 2004. “Hands-on Guide to Questionnaire Research: Selecting, Designing, and Developing Your Questionnaire.” *British Medical Journal*.

### 3.2.3 Questionnaire designing<sup>39</sup>

Writing down a list of questions in a form of questionnaire can be achieved by anyone, but having a good quality data in contrast, as answers, is not always guaranteed.

For having good quality data and reaching efficient and useful responds, the questionnaire should be designed carefully, well-constructed and related to research questions. Otherwise the answers will be inefficient and useless, lead to foggy recommendations and illusive conclusion in addition, which will not enrich the study

**The following steps are essential and should be taken into consideration during the designing of the questionnaire:**

- ❖ Hypothesis identifying<sup>40</sup>: Determine the questionnaire hypothesis that address the research goals and define the required information that should be collected for this study. For instance, is the needed information includes respondents' personal information (age, marital status, ...), professional information (education status, employment status, ...), believes, opinions, daily practices, attitudes.

As a result, to specifying the needed information, the researcher will know which question should be included at the questionnaire.

- ❖ Literature reviewing: revising the previous studies, publications, past and current questionnaire related the subject under study will give the researcher a clearer idea about the conducted study and help him to identifying a clearer direction in his research road map.

This step helps in saving his time and efforts. Furthermore, gaining more knowledge by starting where the others ended and learning from their previous experiences.

---

<sup>39</sup> Roopa, S, and MS Rani. 2012. "Questionnaire Designing for a Survey." *The Journal of Indian Orthodontic Society*.

<sup>40</sup> Roopa, S, and MS Rani. 2012. "Questionnaire Designing for a Survey." *The Journal of Indian Orthodontic Society*.

❖ Questionnaire validity <sup>41</sup>:

It points to the questionnaire accuracy, by ensuring that the questionnaire does measure what it is supposed to be measured. The questionnaire validity is assessed by finding the accuracy of the asked questions and the representativeness of the sample <sup>42</sup>.

❖ Questionnaire Reliability<sup>43</sup>:

It refers to the consistency of the questions in the questionnaire, which bring consistent results from respondents even if the words, questions and statements order and structure are different. Results are similar for repeated samples by different researchers over the time. Differences at the results come from the differences between participants but not from inconsistent understanding of the questions or statement at the questionnaire <sup>44</sup>.

❖ Questions intelligibility and obviousness: Questions must be simple, clear and understandable to all participants. This excludes complex questions. The respondent will answer based on his understanding and in case there is unclear or complex point , this leads to that the participant misunderstands the point and answers based on this misunderstanding. This results in misleading conclusions.

Furthermore, questions must be easy to be answered, visually attractive and free of jargon.

---

<sup>41</sup> Ponto, Julie. 2015. "Understanding and Evaluating Survey Research." *Journal of the advanced practitioner in oncology*.

<sup>42</sup> Boynton, Petra M., and Trisha Greenhalgh. 2004. "Hands-on Guide to Questionnaire Research: Selecting, Designing, and Developing Your Questionnaire." *British Medical Journal*.

<sup>43</sup> Boynton, Petra M., and Trisha Greenhalgh. 2004. "Hands-on Guide to Questionnaire Research: Selecting, Designing, and Developing Your Questionnaire." *British Medical Journal*.

<sup>44</sup> Roopa, S, and MS Rani. 2012. "Questionnaire Designing for a Survey." *The Journal of Indian Orthodontic Society*.

- ❖ Questions accuracy: Direct and short questions are a must here. Each question should contain only one idea at a time for avoiding making the respondent confused by different ideas at the same question. Moreover, questions must be accurate, precise, concert and close to the thinking way of the participants as much as possible.

### 3.2.4 Questionnaire Formulation<sup>45</sup>:

The researcher should focus on the following points when construction the questionnaire<sup>46</sup>:

- ❖ Question sequence: The sequence of questions should move smoothly and in a logic way to enable the respondents to understand the questions clearly and avoid misunderstanding. Questions sequence should be clear as a well to avoid the respondents experiencing a sudden shifting between questions. This raises the participants' desires and encourage them to continue responding to the remaining questions.
- ❖ Questionnaire wording: using simple, clear and easy understandable words as well is a must here. Furthermore, avoiding unclear and ambiguous words and vocabularies. Skipping the dangerous, sensitive and emotional expressions.
- ❖ Questionnaire layout: Questions must be numbered, so the respondents can easily track his path during answering the questions. Avoiding using only capital letters (upper case) words since it is difficult to be read and tiring visually. Classifying the questions into sections, subsections, or groups based on its subject and target. Introducing some instruction for guiding the respondents in answering the questions.
- ❖ Questionnaire Items order: It is an important point and is a complementary step to the questions sequence step. People like to answer the questionnaires in a short time and easily way. The questions order helping the researcher in giving the respondents

---

<sup>45</sup> Song, Youngshin, Youn Jung Son, and Doonam Oh. 2015. "Methodological Issues in Questionnaire Design." Journal of Korean Academy of Nursing.

<sup>46</sup> Spangenberg, Eric R., Frederick T. L. Leong, and James T. Austin. 1998. "The Psychology Research Handbook: A Guide for Graduate Students and Research Assistants." Journal of Marketing Research.



this impression. Starting with the item or question that is related directly to the questionnaire goal and purpose. The rest of the items are ordered based on the targeted respondents.

### 3.3 Sampling

The research method is designed and formatted to collect data from people, but asking every single individual from the whole population by the research method (which is here the questionnaire) is, unachievable, impractical, very expensive and time consuming.

The sampling strategy here is the solution, it can be achieved by targeting a sample of people from the whole population that is under study. The sample should have respondents with similar characteristics as the whole population

The sample should be sufficient and representative to the population of interest. Recruiting the sample participant in different way is a guarantee for having larger sample. As much as the sample is random and large, the participated respondents reflect the main features of the whole population of interest. By this way, the sample strategy will lead to accurate conclusions and better understanding to the population that is under study. Sampling is classified under two main categories:

- ❖ Random sampling: It is employed at the case of collecting quantitative data by using quantitative data collection method (for instance questionnaire). Statistical analysis is implemented here if it is appropriate. The data is collected randomly to represent the whole population of interest. This category includes the following main sampling techniques: simple sampling, cluster sampling, systematic sampling and stratified.
- ❖ Non-random sampling: It is employed to collect qualitative data using the qualitative data collection methods, such like: focus groups and interviews. The target is to achieve exploratory work. The main sampling techniques here are : Snowballing, Convenience sampling, purposive sampling:

Two last important points should be taken into consideration during the sampling process, which are:

**Sample size:** generally, as much as the sample is large with suitable sampling selection technique, reaching accurate results and conclusions is more achievable. But practically this is not always easy as basically it is expensive, efforts and time consuming. Scholars have agreed that the sample size for a survey research depends on three main factors: resources availability, the conducted study goals and purposes

**Sampling error:** sampling error occurs when the sample does not represent the whole population. It can't be eliminated completely, but its extend can be controlled by sampling techniques.

### 3.4 Piloting

Before implementing the questionnaire and sending it through post or online, there is a preliminary step is called Piloting. It is a preparatory step where the researchers choose a small group of individuals from the planned targeted sample to the conducted questionnaire. The main goal of this step is to check the effectiveness of the questionnaire, the questionnaire pilot study will check the following:

- ❖ If the questions are ordered in the best way or are in a need to be reordered?
- ❖ Are the questions understood well and easy to follow to all respondents?
- ❖ Are wording, expressions and instructions clear to all respondents or is there any misunderstanding?
- ❖ Is there any specific question, instruction or comment need to be added?
- ❖ Are the questions framed in a way to reach the desired results?

The notes, conclusions and recommendations from this step should be taken into consideration by the researchers. Any problems or difficulties that discovered during holding the pilot study should be revised by researchers before implementing the questionnaire. This step mainly aims to improve the questionnaire design, formulating and formatting before implementing it.

### **3.5 Questionnaire analysis**

The purpose of this step is to explore the data, study it, understand its attributes and find the relations between it. It is the core step at the survey research as it summarizes the collected data, gives conclusions and answers the research questions. The researchers should put into their consideration that this step needs its enough time since if it has been achieved quickly, important data details could be missed and wrong analysis could be conducted, this lead to inadequate conclusions.

The used analysis method is chosen based on the research method and the survey research questions. This step should be considered carefully at the planning steps of both the research method and the survey research.

# Chapter 4

## Political participation of immigrant women in Norway during the local elections in 2015 and the national elections 2017 questionnaire

### Summary

This chapter is divided to two sections; The first section presents the goals of the conducted research, the selected research method, the aim of the constructed questionnaire, the targeted group of the questionnaire and the questionnaire hypothesis. In additions, it reviews some of the previous literature about women's political participation.

The second section discuss the questionnaire construction . It presents the design, formulation, formatting , administration and spreading of the questionnaire.

### 4.1 The constructed questionnaire

#### 4.1.1 The research goals

The main goal of this study is to learn more about immigrant women's political participation in Norway during the local elections 2015 and the national elections 2017. The data collection will be achieved through the questionnaire we have built. While the data analysis will be implemented by using the statistical methods; the principle component analysis, binomial multiple logistic regression and the principle components logistic regression( are presented in details in chapter 6)

#### 4.1.2 The selected research method

By reviewing the mentioned questions in Section 3.2.1, about choosing the data collection tool, furthermore asking the following questions<sup>47</sup>:

- ❖ Is the questionnaire the best way to collect the information for this research?  
Is it the most suitable method for the situation? why?
- ❖ Will the questionnaire add to the data analysis? What? How?
- ❖ Will the questionnaire address the research problem and contribute in finding answers to the research question? How?
- ❖ Will the questionnaire be formulated to reach the point at the previous question? How?
- ❖ Is it clear to the researcher who are the targeted respondents to this research methodology?
- ❖ Will the researcher be able to reach the respondents? How?
- ❖ Will the respondents have clear and enough information enable them from answering the questionnaire? How?
- ❖ Will the questions be clear enough to the targeted group?
- ❖ Will the targeted group be willing to participate and answer?

And Based on that the answer for the polar questions<sup>48</sup> (Yes\No) were all yes, moreover since the planned collected data is quantitative, the selected research method to this study here is the ONLINE QUESTIONNAIRE <sup>49</sup> .

---

<sup>47</sup> Laaksonen, Seppo. 2018. *Survey Methodology and Missing Data: Tools and Techniques for Practitioners* *Survey Methodology and Missing Data: Tools and Techniques for Practitioners*

<sup>48</sup> Colosi, Laura. 2006. Cornell Cooperative Extension Designing an Effective Questionnaire.

<sup>49</sup> Ball, Helen L. 2019. "Conducting Online Surveys." *Journal of Human Lactation*.

### **4.1.3 The questionnaire aim**

Studying the immigrant women's political participation in Norway during the local elections 2015 and the national elections 2017, by finding the main factors that affected their participation in the ballots in specific and their political participation in general in addition to mapping the relation between these variables.

### **4.1.4 Targeted group of the questionnaire**

Immigrant women in Norway.

### **4.1.5 Literature and previous work reviewing**

What do we know about immigrant women in general and their political participation in specific is so little. There is not enough segregated data about them in the majority of the host countries. The data at their new countries are usually divided based on gender (women and men) or based on the background (immigrant and non-immigrant). But it isn't always based on these 2 factors together (gender and immigration background). We do need to know more about their integration status and what are the obstacles which confront the integration process. The best way to measure their integration progress is to measure the economic, social and political integration in the society <sup>50</sup>.

For immigrants (women and men) in general and immigrant women in specific, their political participation plays an important role in their overall integration process. It is an indicator that shows how much are they connected to the society. Since it is a tool to express their interests, thoughts, opinions and believes about the policies, politicians and the whole political system at the host country <sup>51</sup>.

Based on the previous studies and literature in this field, which is not that much by the way to the moment, the main restrictions and challenges which confronts the immigrant women's

---

<sup>50</sup> O'Neill, Brenda., Lisa; Young, Elisabeth; Gidengil, and Catherine Côté. 2013. "The Political Integration of Immigrant and Visible Minority Women." *Canadian Political Science Review*.

<sup>51</sup> Bermudez, Anastasia. 2016. *International Migration, Transnational Politics and Conflict. The Gendered Experiences of Colombian Migrants in Europe*.

political participation are; learning new language at the host country which is neither easy nor happens quickly, the traditional and cultural values they respect and follow, the limitation of their economic and social integration at the new country <sup>52</sup>.

After several investigations to find more about the causes and reasons to the current political participation status, new factors showed up that shape their political participation such as <sup>53</sup>; they are isolated since the large families they are responsible about and overloaded with them and with the house cohort, they don't trust politics or the political system based on an experience they keep in their memories from the country of origin, they don't believe that they could make a change at their new countries, they don't have enough time to participate in politics, politics is not at the top of their priorities' list or even they don't know how to navigate the political system in order to the lack of information they have <sup>54</sup> .

#### **4.1.6 Questionnaire building**

Using a previous questionnaire or taking some questions from previously existing questionnaire, increases the validity and reliability<sup>55</sup> of the new modified questionnaire; but this is only in the case if it is used for the same goals and aims that it was built for. It can be used for related goals but at this case it should be modified to suit the required goals <sup>56</sup>.

Validity and reliability increase here, because there is an implicit guarantee that these previous questions are clear to all the previous targeted respondents. The questions are easy to follow, wording and expressions are understandable by all respondents and the framed

---

<sup>52</sup> O'Neill, Brenda.; Lisa; Young, Elisabeth; Gidengil, and Catherine Côté. 2013. "The Political Integration of Immigrant and Visible Minority Women." *Canadian Political Science Review*.

<sup>53</sup> Bermudez, Anastasia. 2016. *International Migration, Transnational Politics and Conflict. The Gendered Experiences of Colombian Migrants in Europe*.

<sup>54</sup> Bilodeau, Antoine. 2016. "Migrating Gender Inequalities? Immigrant Women's Participation in Political Survey Research." *International Migration Review*.

<sup>55</sup> Roopa, S, and MS Rani. 2012. "Questionnaire Designing for a Survey." *The Journal of Indian Orthodontic Society*.

<sup>56</sup> Song, Youngshin, Youn Jung Son, and Doonam Oh. 2015. "Methodological Issues in Questionnaire Design." *Journal of Korean Academy of Nursing*.

questions bring the expected results. Moreover, the pilot study phase is shorthanded at the case of using previous questions or questionnaires. There is no need to pilot the questionnaire and check that the asked questions are simple and understood, since this was done practically before<sup>57,58</sup>

For increasing the validity and reliability of our questionnaire and based on the shortening in the available time which we have to use to achieve this study, plus to avoid misleading conclusions in such a study in order to the limited quantitative previous study in this field, we have decided to use previous questionnaires from inside and outside Norway. Consequently, the piloting study is not needed.

Our questionnaire was constructed by selecting the needed questions from previous related questionnaires and studies. Some of these questions were taken as it is. Moreover some questions were modified slightly to adapt our needs. Furthermore, we have formulated few extra questions.

The main resources here and the previous questionnaires and studies that are used in building the questionnaire are;

- ❖ Velgerundersøkelse blant personer med innvandrerbakgrunn 2013 Dokumentasjonsrapport, SSB<sup>59</sup>.
- ❖ Ballots and belonging- new citizens on political participation -full report, ICC insights<sup>60</sup>.
- ❖ The Political Integration of Immigrant and Visible Minority Women<sup>61</sup>.
- ❖ Residency formalities by European Union<sup>62</sup>.

---

<sup>57</sup> Institute Canadian Citizenship. 2015. BALLOTS & BELONGING-NEW CITIZENS ON POLITICAL PARTICIPATION.

<sup>58</sup> Song, Youngshin, Youn Jung Son, and Doonam Oh. 2015. "Methodological Issues in Questionnaire Design." Journal of Korean Academy of Nursing

<sup>59</sup> Høstmark, Maria-SSB. 2013. Velgerundersøkelse Blant Personer Med Innvandrerbakgrunn 2013.

<sup>60</sup> Institute Canadian Citizenship. 2015. BALLOTS & BELONGING-NEW CITIZENS ON POLITICAL PARTICIPATION.

<sup>61</sup> O'Neill, Brenda;, Lisa; Young, Elisabeth; Gidengil, and Catherine Côté. 2013. "The Political Integration of Immigrant and Visible Minority Women." Canadian Political Science Review.

<sup>62</sup> European Union. "Residency Formalities by European Union. -Permanent Residence (>5 Years) for EU Nationals." 2018.



- ❖ The European Social Survey (ESS)ESS8 edition 1.0 <sup>63</sup>.
- ❖ Canadian Election Study 2015 <sup>64</sup>
- ❖ Aventure Linguistique-your language level.

#### **4.1.7 The questionnaire hypothesis identification:**

The two main questions of the questionnaire are:

- 1) Did you vote in the 2017 national elections?
- 2) Did you vote in the 2015 local elections?

These questions are the responses or the dependent variables. The main goal here is to find the linear relation between each dependent variable aside with the other explanatory variable. In addition, finding the relation between the independent variables which have significant relation with the dependent variables.

## **4.2 Questionnaire establishment**

### **4.2.1 Questionnaire design**

The questionnaire is available in English and Norwegian languages. The participant can choose which language she would like use, but she can only answer one time. This point is clarified at the introduction of the questionnaire.

Furthermore, there is an introduction that clarifies the aims and goals of the questionnaire and that it is a part of a master thesis in the data analysis master program at the University of Bergen. The introduction is written in English at the first page but if the participant chooses the Norwegian introduction, she will find it in Norwegian. The contact information for the questionnaire responsible persons are given in the introduction as well.

---

<sup>63</sup> European Social Survey. 2016. ESS8-2016, Ed.1.0 Study Documentation.

<sup>64</sup> Fournier, Patrick, Fred Cutler, Stuart Soroka, and Dietlind Stolle. Canadian Election Study 2015.

There is a confirmation that the participation will be anonym. And it is mentioned that it is voluntary to fill in this questionnaire, and the participants can at any time decide to stop filling in the questionnaire without mentioning any reason. She simply can close the window without submitting.

There is a confirmation that this questionnaire is constructed and developed based on previous questionnaires and call interviews provided by similar studies. In addition, it is mentioned that the study has been notified to the Data Protection Official for Research, NSD - Norwegian Centre for Research Data.

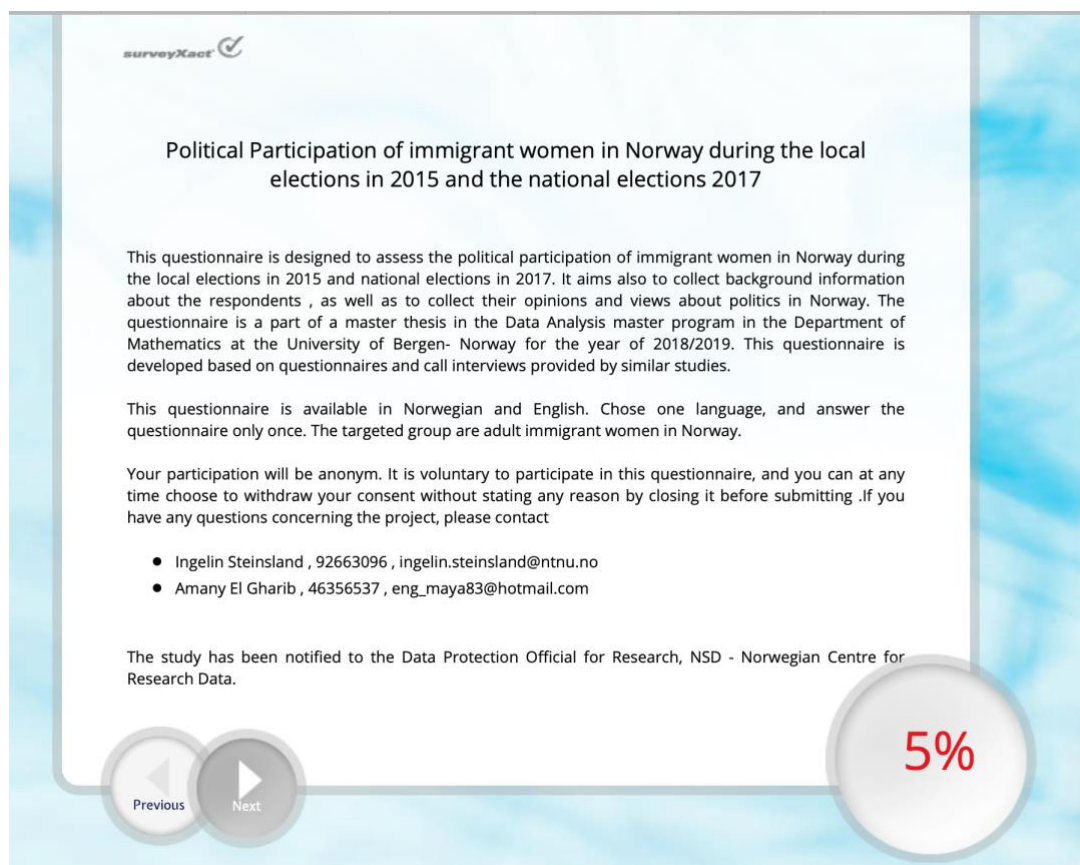


Figure (4.1): The front page of the established questionnaire

Once the respondent press Next, she will be directed to the second main page where she should confirm receiving enough information about the project and she is willing to participate in filling the questionnaire. She will never be directed to the next page without

pressing yes, and then press Next button. The respondents can choose the language of the questionnaire. In case she selects Norwegian she will have the previous two pages again but in Norwegian. And this is to be sure that she has the enough information about the project. At the two versions, the followed page is a confirmation from the participant that she is an immigrant woman based on the SSB definition( *immigrant woman : is the woman who immigrated herself to Norway or who was born in Norway for two foreign-born parents*)<sup>65</sup> and she is over 18 years old. Since the targeted group is immigrant women in Norway. The respondents will never be directed to the next page without selecting “Yes” for both questions and then press the “Next” button

In general, the participant has to answer all the questions that appear to her. Otherwise she will be stuck at the page where she stops, and she will not be directed to the next page without answering all questions at her current page.

At the end of the questionnaire there is a submit button and a “Thank you “message to the participants.

#### **4.2.2 Questionnaire formulation and formatting**

The questionnaire consists of the following five main sections:

- ❖ Personal information
- ❖ Information about the residency in Norway
- ❖ The Political Participation
- ❖ The surrounded environment and your political view
- ❖ The Political View

The first section summarises a general overview about the participant’s personal information; age, social status, educational level, employment status, family main income, how many family members the participant responsible about and the degree of sharing the household chores.

---

<sup>65</sup> Byberg, Ingvild. 2002. Immigrant Women in Norway A Summary of Findings on Demography, Education, Labour and Income.

For avoiding the sensitive information, some questions were categorised in groups instead of giving exact answer. For instance

Age:

- 18- 24
- 25- 44
- 45-66
- 67- 79
- 80 or above.

The second section gathers information about the country of origin and residency in Norway. These questions are; the home of origin, the suitable status describe her residency in Norway ( she was born in Norway or she moved to Norway), the period of the residency in Norway, the type of her residency permit in Norway, the current address, the Norwegian language level. The “current address” question is categorised to the regions in Norway and not the towns to avoid gathering sensitive information here as well.

Based on some of the previous studies, the most important indicator here is the language which is a big challenge to several women who immigrated to Norway<sup>66</sup>. The type and period of the residency play an important role which affect the language indicator since the new immigrant in general have more challenges in learning the language more than people who immigrated to Norway long time ago. People who are citizens or holding permanent residency at the country theoretically have more motivation to learn the language than the people who have asylum seeking status, or student visa for example.<sup>67</sup>

The third section discusses the respondent’s political participation. It contains the following questions; are you interested in politics ?did you vote in 2017 national elections ?did you vote in 2015 local elections? why did you vote ?why didn’t vote? did you vote at the in her home country, in case the ?If you have moved to Norway, did you vote at your home country elections before moving to Norway? have you been a member of a political party, at the past 12 months? have you been a member of an interest group working for a change on a

---

<sup>66</sup> O’Neill, Brenda;, Lisa; Young, Elisabeth; Gidengil, and Catherine Côté. 2013. “The Political Integration of Immigrant and Visible Minority Women.” *Canadian Political Science Review*.

<sup>67</sup> Institute Canadian Citizenship. 2015. *BALLOTS & BELONGING-NEW CITIZENS ON POLITICAL PARTICIPATION*.

particular social or political issue, at the past 12 months? have you been a volunteer for a group or organization other than a political party, at the past 12 months ? have you taken a part in a demonstration, at the past 12 months? have you participated in signing a petition, at the past 12 months?

This section will examine the participant's capacity and ability to navigate the political system in Norway . One of the main reasons to not participate could be the lack of knowledge about the voting system, the voting campaigns, candidates and parties <sup>68</sup>. On the other hand, the barriers that confront the respondent's participation could be that; she doesn't believe that her participation is important, or she believes that the raised issues aren't in her own interests <sup>69</sup>. The language is a big obstacle which prevents her from voting. She doesn't have enough time to vote, could be a possible reason. Some people who came originally from countries that had corrupted and dictatorial governments or regimes don't trust politics in general. Simply they don't vote or sometime are scared to vote.

On the other side the motivation of the people who did vote could be simply their desire to make a change by making their voices heard <sup>70</sup>. They could do this change by supporting a particular party or a particular candidate so they voted because of this<sup>71</sup> . They are interested to support a particular raised issue at the campaign, is a possible reason. Other people believe that they should vote as long as they are eligible "citizens or holding the permanent residency". Other reason to vote is that the family or friends persuaded the participant to vote <sup>72</sup>.

---

<sup>68</sup> O'Neill, Brenda;, Lisa; Young, Elisabeth; Gidengil, and Catherine Côté. 2013. "The Political Integration of Immigrant and Visible Minority Women." Canadian Political Science Review.

<sup>69</sup> Community relations, and Directorate of Social and Economic Affairs Council of Europe Press. 1995. IMMIGRANT WOMEN AND INTEGRATION.

<sup>70</sup> Wong, Winnie, and Yves Poisson. 2008. FROM IMMIGRATION TO PARTICIPATION: A Report on Promising Practices in Integration.

<sup>71</sup> Bilodeau, Antoine. 2016. "Migrating Gender Inequalities? Immigrant Women's Participation in Political Survey Research." International Migration Review.

<sup>72</sup> Institute Canadian Citizenship. 2015. BALLOTS & BELONGING-NEW CITIZENS ON POLITICAL PARTICIPATION.

The fourth section is about the surrounded environment and the participants' political view and thoughts. The questions are; how often did the family discuss politics at the past?, how often does the participant discuss politics now?, how many immigrant friends does the participant has? and to which degree does family or friends affect in choosing the political party that she supports?

The fifth section measures the respondent's political interests and believes, for example; to which degree does the participant trust politics, political parties and the government in Norway? to which degree does the participant believe or having interest in making a change by voting? to which degree is the participant engaged in politics in Norway? Should the politicians focus on the immigrant women's issues in Norway? to which degree is the participant satisfied with the Norwegian political system? to which degree does she understand the voting system? to which degree does traditions affects her political participation? is she financially independent? Does she concern herself as a strong supporter to women's rights?

#### **4.2.3 Questionnaire administration and spreading**

The questionnaire is online. Since it collects valuable information, we have to use a service provider which has an agreement with university of Bergen.

The questionnaire was designed, programmed, published and managed through Surveyxact. Which is an internet based system for designing and managing surveys. A comprehensive user manual for this service provider , in English:

<https://view.publitas.com/ramboll/surveyxact-user-manual-12-5/page/1>

The questionnaire was self-administrated by the researcher. It was sent to several different non-governmental organizations, municipalities, governmental organizations, political parties and International organizations by sending the direct hyperlink by e-mail.

---

The questionnaire was published as well at the social media pages. Some live sessions were hold on Facebook to translate it live from Norwegian language to Arabic language based on Arabic women request.

#### **4.2.4 Sampling**

The questionnaire distributed more than 2000 times. 1816 didn't open it or open the front page and didn't continue. While the complete responds are 95. The partially completed cases are 95, and the rejected cases 27.

Again the targeted places that received the questionnaire links were, several NGOs and INGOs in Norway, political parties, universities, governmental organizations. It was spread as well to the social media pages and groups that target immigrants generally, women in N

# Chapter 5

## Data and Explanatory analysis

### Summary

This chapter has eight sections; the first three sections previews the data preparation, the whole data general description and the variables preparation. The following five sections describes each data set from the five data set aside.

### 5.1 Data preparation

The survey was distributed by links, and 1816 links were distributed but the receiver either opened only the front page without any interaction and then close it or didn't open the link at all. The total number of the participants who opened the questionnaire are. And these are divided into three levels; rejected (27 participants), partially completed (95 participants) and completed (95 participants). Participants were categorized as rejected if they are not in the target group that; above 18 years old and immigrant and woman.

Almost all the participants who partially completed the questionnaire, have stopped answering before the key questions about voting in the elections in 2015 and 2017. So the partial responds were excluded from the data analysis. Only the 95 completed records are used to implement the analysis.

### 5.2 The collected data description

The dataset contains 52 explanatory variables plus two dependent variables. There are 4 categorical variables:

- Age
- Education level
- Language level
- Social Status



Despite that the boxplot is a primitive plot in comparison to histogram and density plot for example, but in less space it gives first impression about the collected data and the data variables. It shows the approximate minimum, maximum and the median values. It shows if there is outlier values as well.

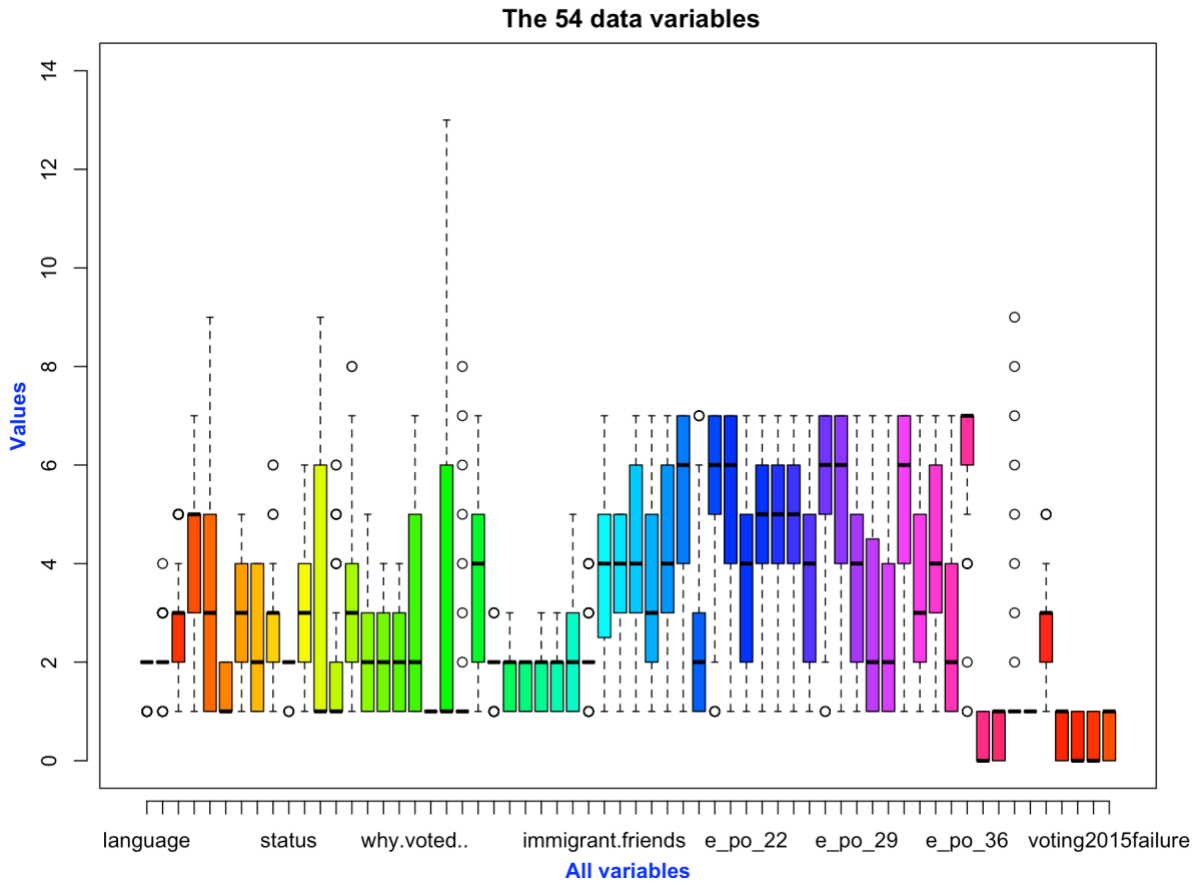


Figure (5.1): The collected data 54 variables

Figure (5.1) shows boxplot of the data variables. For example it shows that the maximum value for the " why did vote" variable =13. While the variable's " voting 2015 failure" values are 0 and 1. The boxplot doesn't cover all the 54 explanatory variables here but one can find illustrative plot that describes the values to some explanatory variables.

Since the questionnaire design has five main sections, the collected data are divided into five main questionnaire classes data sets : personal information, information about the residency in Norway, the political participation, the surrounded environment and your political view

and the political view. These classes will be called from this point to the end of the thesis as the five data sets ; “personal information” data set, “information about the residency in Norway” data set, “the Political Participation” data set, “the surrounded environment and your political view” data set, and the “political view” data set.

### **5.3 Variables preparation**

The two dependent variables at this conducted study are: "Voting in 2015 local elections" and "Voting in 2017 national elections". Both of them have 4 values, which are: “yes=1, no=2, not eligible=3 and refuse to answer=4”. To be able to use them in the logistic regression analysis, “3” and “4” values were converted to NA. Then converted the “2” value to 0. So it have 0 or 1 values which is required to the binary response for the binomial multiple logistic regression.(more details in chapter7)

There are three questions that appear based on the participant’s answer, the questions are respectively; “If your voted in one or both 2015,2017 elections. Why did you vote?”, “If you didn’t vote in one or both 2015, 2017 elections Why didn't you voted?” and “Did you vote in your country at any elections?”. Preparation of these variables is mentioned in chapter 6 The first question only appears if the respondent answered one or both of the dependent variables: "Voting in 2015 local elections" and "Voting in 2017 national elections" with “Yes”. The second question only appears if the respondent answered one or both of the dependent variables: "Voting in 2015 local elections" and "Voting in 2017 national elections" with “No”. The third question only appears if the respondent answered the question “Which of the following does describe your status better? status is “with “I have moved to Norway”.

### **5.4 Personal information data set (x1)**

This is the first data set, and it is referred to it by (x1). It contains seven personal information variables that are; age, social status, education status, employment status, who has the main income in the respondent’s family, the family members number in the respondent’s household and she is responsible for taking care of (e.g. children, elderly relatives...) and the division of household chores between she and her partner.

### 5.4.1 Personal information variables general description

- Age: the “Age” variable is classified to five levels, as follows:
  - ❖ Level 1: from 18 to 24 years old.
  - ❖ Level 2: from 25 to 44 years old.
  - ❖ Level 3: from 45 to 66 years old.
  - ❖ Level 4: from 67 to 79 years old.
  - ❖ Level 5: 80 and above years old.

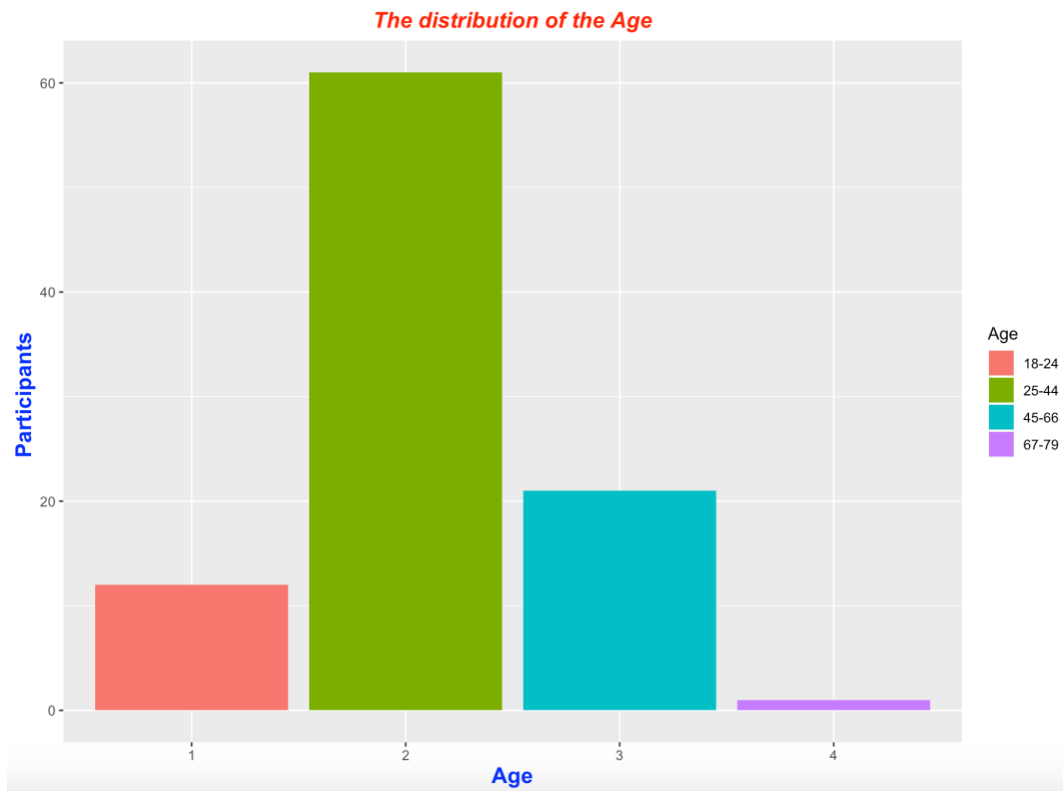
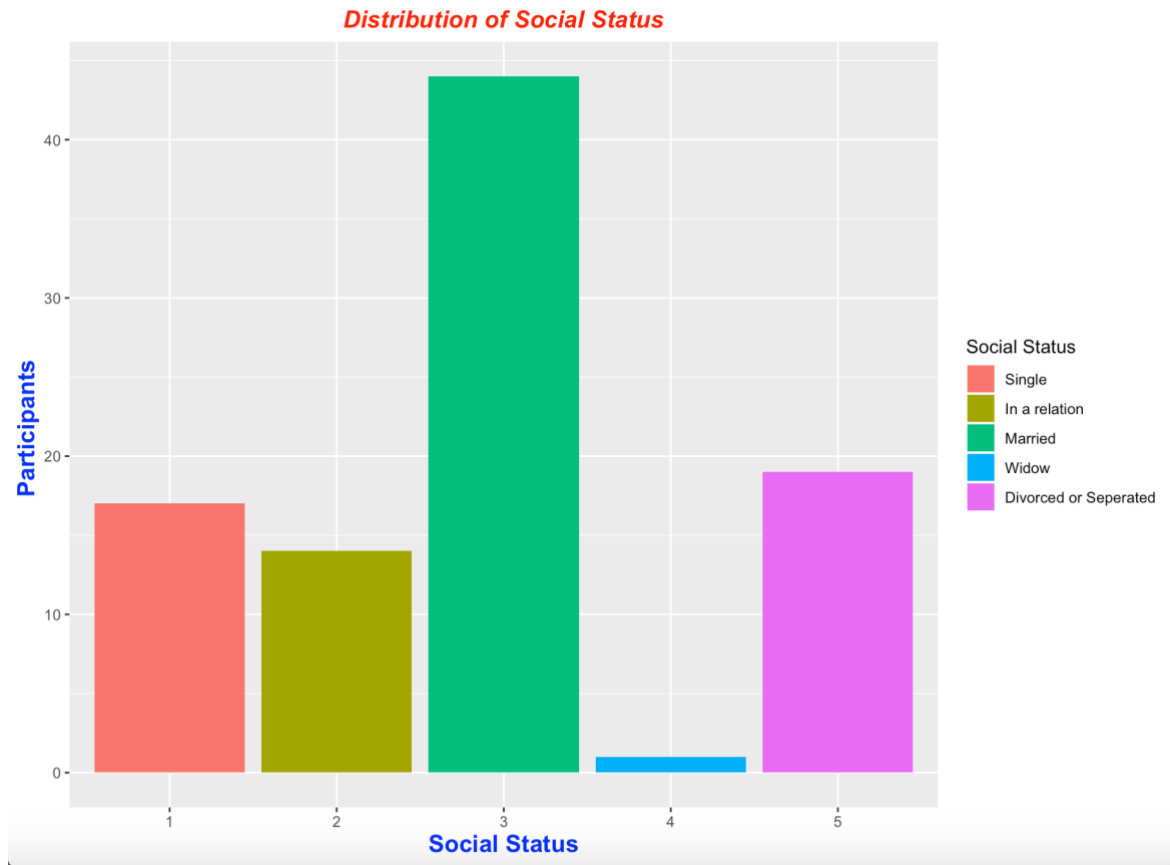


Figure (5.2) : The “Age” variable distribution among the participants

Figure (5.2) shows that the majority of the participants’ age locates in level (2) and level (3). Which means that the majority of respondent to this questionnaire are from 25 to 44 years old. No participant from the last lever (80 or above)

- Social Status; the “ Social status” variable is classified to five levels, as follows:
  - ❖ Level 1: Single
  - ❖ Level 2: In a relationship.

- ❖ Level 3: Married
- ❖ Level 4: Widow
- ❖ Level 5: Divorced or Separated



Figure(5.3) The "Social status" variable distribution among the participants

Figure (5.3) shows that the social status variable distribution from level 1 to level 5. The majority of the participants here are married level (3), then divorced or separated level (5).

- Educational status; the "educational status" variable is classified to seven levels, as follows:
  - ❖ Level 1: Didn't complete secondary school
  - ❖ Level 2: Completed Secondary school
  - ❖ Level 3: Completed High school
  - ❖ Level 4: Completed College or technical school diploma
  - ❖ Level 5: Completed University - Bachelor's
  - ❖ Level 6: Completed University - Master's

❖ Level 7: Completed University – PhD

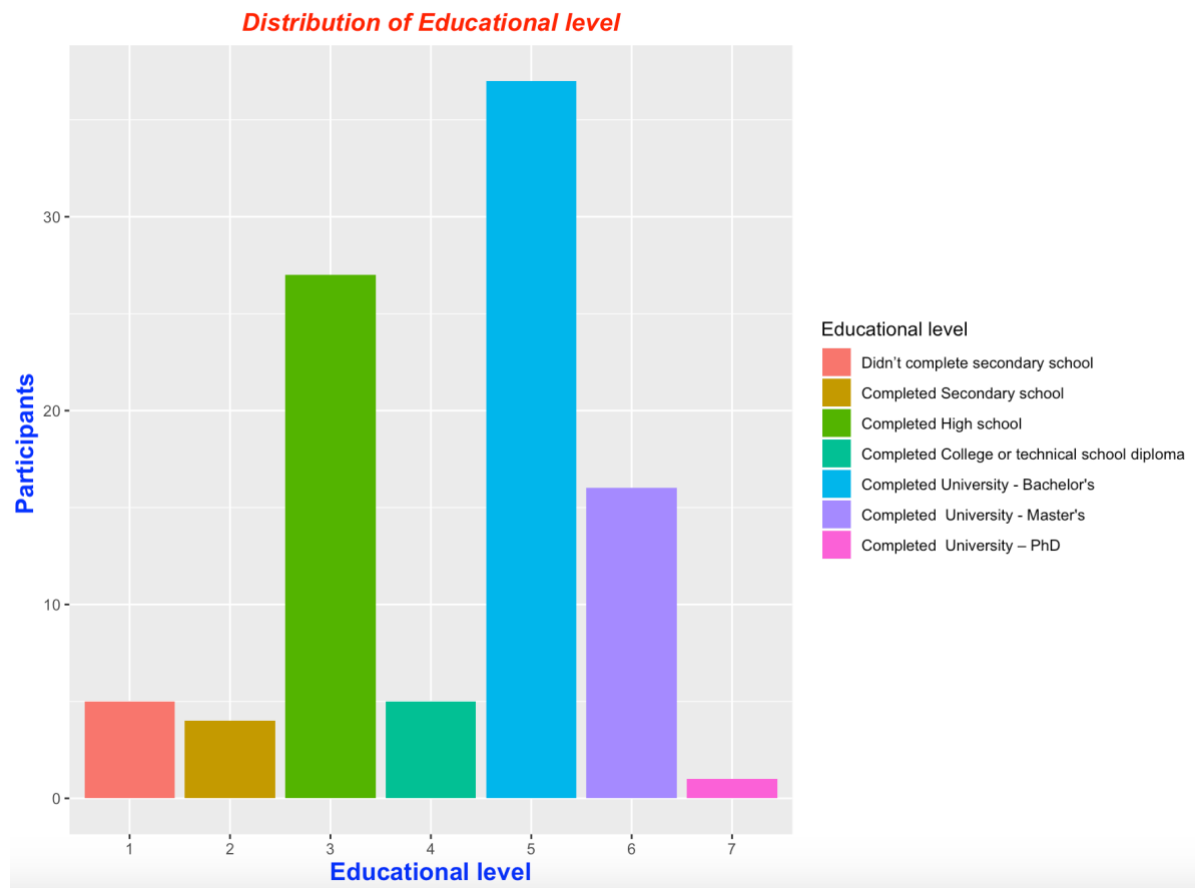


Figure (5.4): The “Educational status” variable distribution among the participants

Figure (5.4) shows that the educational status variable distributes from level 1 to level 7. The majority of the participants’ educational status locates in level (5) then in level (3), then in level (6). Which indicates that the majority of the respondents to the questionnaire completed their Bachelor’s degree.

- Employment status; the “ Employment status” variable is classified to nine levels, as follows:
  - ❖ Level 1: Employed full time
  - ❖ Level 2: Employed part time
  - ❖ Level 3: Self-employed
  - ❖ Level 4: Searching for work (unemployed)

- ❖ Level 5: Caring for family members at home (e.g. children, elderly relatives)
- ❖ Level 6: Retired
- ❖ Level 7: Student
- ❖ Level 8: Working and attending school
- ❖ Level 9: Other

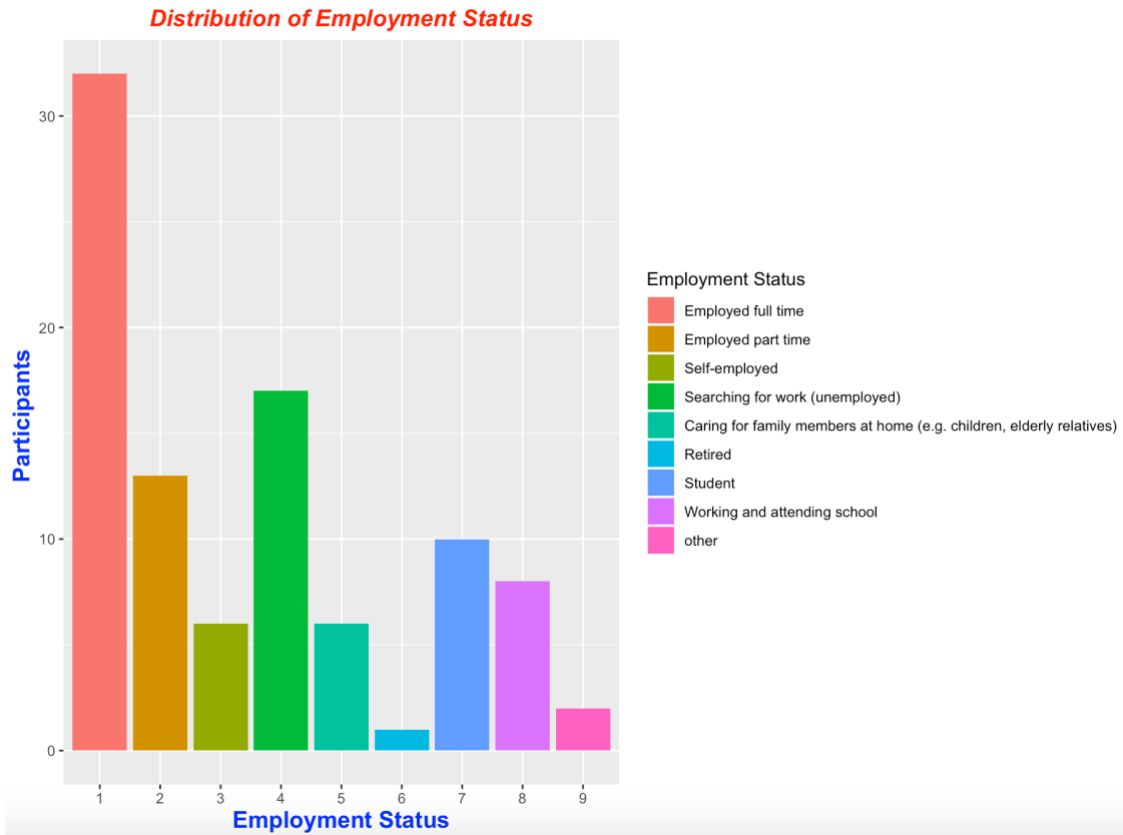
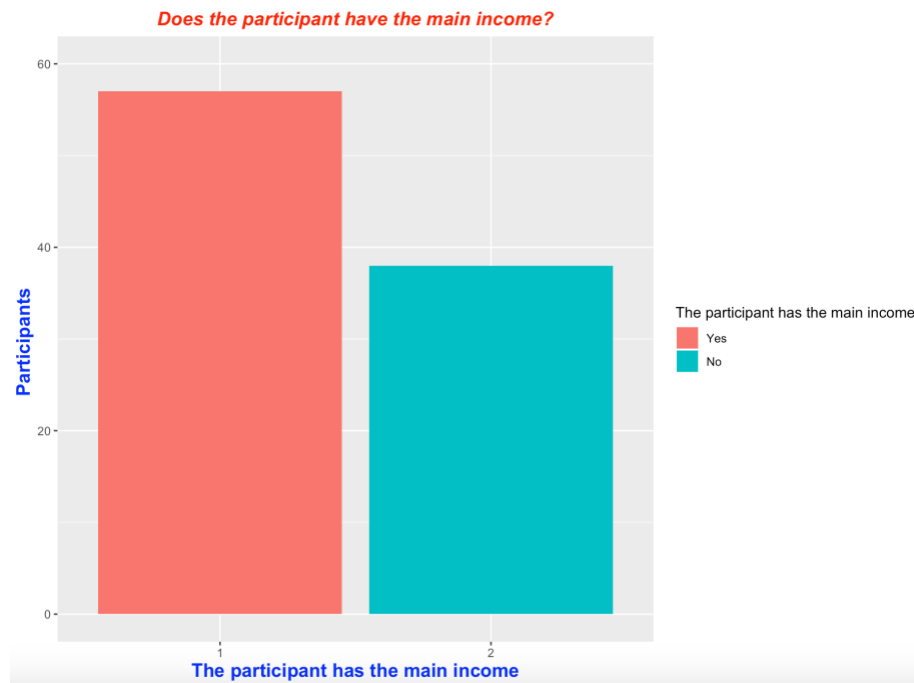


Figure (5.5): The “Employment status” variable distribution among the participants

Figure (5.5) shows that the employment status variable distribution among the participants. The majority of the participants’ employment status locates in level (1) “employed full time”, then in level (4) “completed college or technical school diploma”. This indicates that the majority of the participants are employed full time.

- The main income of the family; the “Main income of the family” variable asks the participant if she has the main income of the family or not. It is classified to two levels, as follows:
  - ❖ Level 1: Yes
  - ❖ Level 2: No



Figure(5.6) : The “main income of the family” variable distribution among the participants

Figure (5.6) shows the main income in the family variable distribution. And it indicates that the majority of the participants have the main income in the family.

- The family members number the respondent is taking care of; is classified to five levels, as follows:
  - ❖ Level 1: None; I’m responsible only for myself
  - ❖ Level 2: One
  - ❖ Level 3: Two
  - ❖ Level 4: 3-5
  - ❖ Level 5: More than five

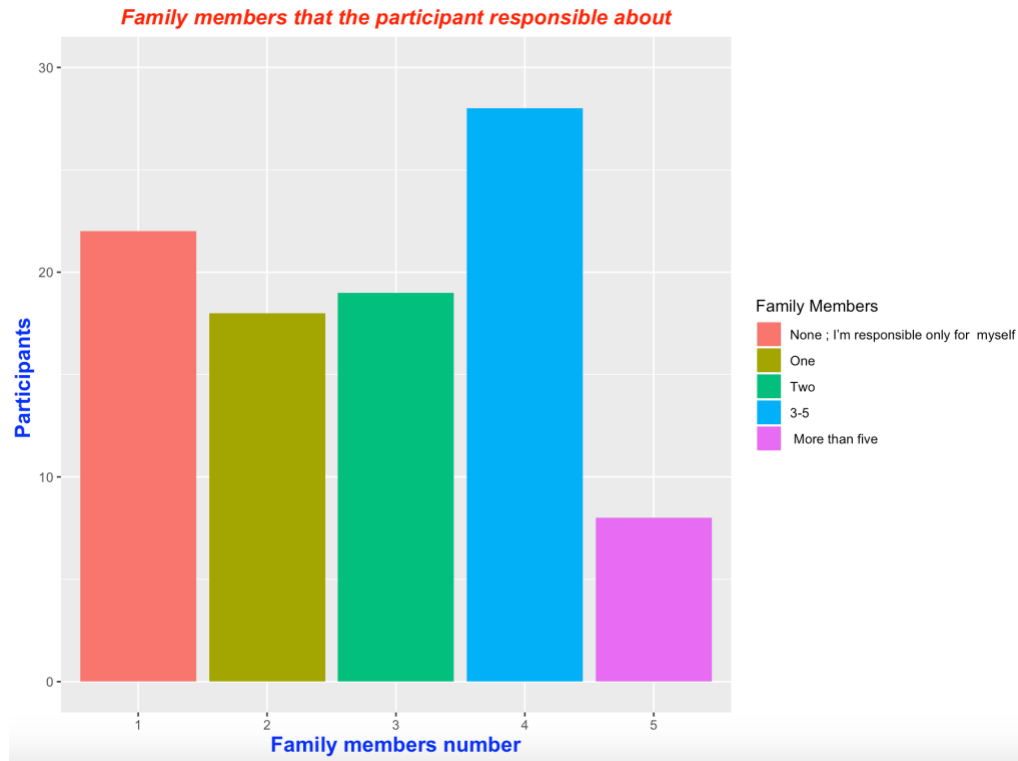


Figure (5.7): The “family members number” variable distribution among participants

Figure (5.7) shows “the family members number that the respondent is taking care of” variable distribution from Level (1) to Level (5). majority of the participants are taking care of 3-5 participants (Level 4)

- The division of household chores; is classified to four levels, as follows:
  - ❖ Level 1: I don’t live with a partner
  - ❖ Level 2: I do most of the chores.
  - ❖ Level 3: My spouse/partner do most of the chores.
  - ❖ Level 4: We share the chores equally



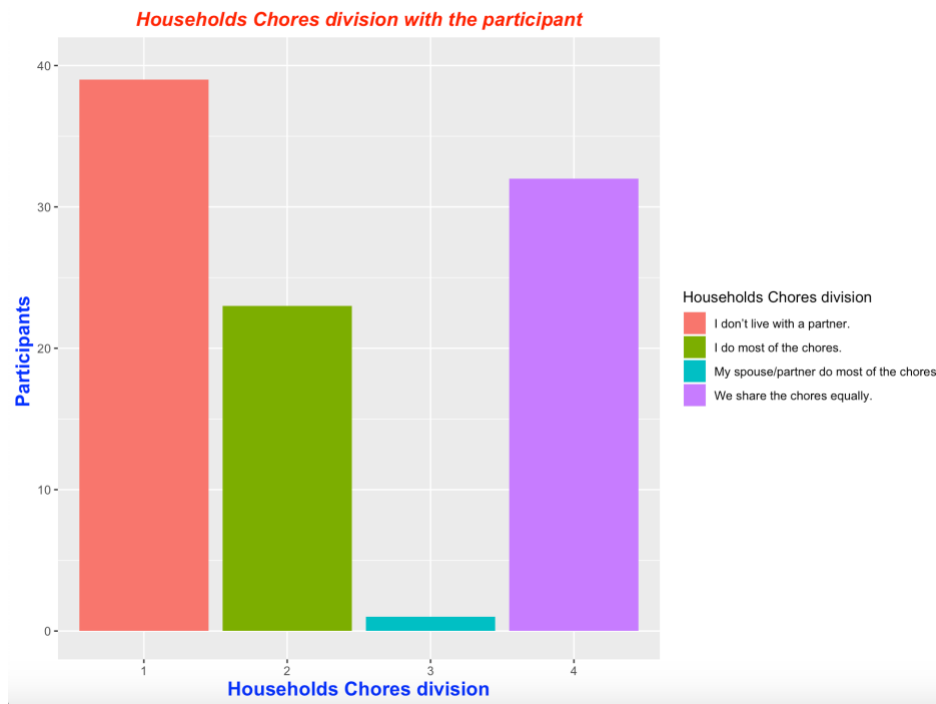


Figure (5.8): The “division of household chores” variable distribution among participants

Figure (5.8) shows that the majority of the participants have chosen level (1), then level (4).

## 5.4.2 Correlation

The correlation is defined always by the linear relationship between two quantitative variables <sup>73</sup>. When the increasing of one variable follows to increasing to the other variable, the correlation is known here by positive correlation <sup>74</sup>. But when the increasing of one variable follows to decreasing to the other variable, the correlation is known by negative correlation.<sup>75</sup>

---

<sup>73</sup> Akoglu, Haldun. 2018. “User’s Guide to Correlation Coefficients.” *Turkish Journal of Emergency Medicine*.

<sup>74</sup> Mukaka, M. M. 2012. “Statistics Corner: A Guide to Appropriate Use of Correlation Coefficient in Medical Research.” *Malawi Medical Journal*.

<sup>75</sup> Pak, Son Il, and Tae Ho Oh. 2010. “Correlation and Simple Linear Regression.” *Journal of Veterinary Clinics*.

By finding the correlation matrix between the personal information variables, some of the following results are found:

- ❖ Educational level increases with increasing of age
- ❖ Main income increases with better job situation.
- ❖ Household chores increases with increasing of family member

### **5.4.3 Collinearity and multicollinearity**

The problem for the collinearity and multicollinearity that it affects the regression estimates and follow to high standard errors. The multicollinearity can be defined here by finding the eigenvalues and eigenvectors. Then check the eigenvalues after decomposing them from the `eigen()` function and check ; the decreasing order ,the maximum and the minimum values and the ratio between the maximum and minimum eigenvalues . If the ratio is less than 100 so the multicollinearity is not a big issue here while if it is more than 100 that means that the multicollinearity problem here is significant and we need to fix.<sup>76</sup>

For x1 dataset's explanatory variables, the ration between the maximum and minimum eigenvalues = 5.07. This indicates that the multicollinearity is not significant issue between the x1 data set's explanatory variables.

### **5.4.4 Correlation relation between personal information variables and voting in 2017 national elections variable**

The highest positive correlation here is between “voting 2017 national election” variable and education status=0.23.

Probability of voting in 2017 increases with higher education and having the main income.

---

<sup>76</sup> Dormann, Carsten F. et al. 2013. “Collinearity: A Review of Methods to Deal with It and a Simulation Study Evaluating Their Performance.” *Ecography*.

#### **5.4.5 Correlation relation between personal information variables and voting in 2015 local elections variables**

The highest positive correlation here is between “voting in 2015 local elections” variable and age =0.30. The highest negative correlation is with employment status variable= - 0.22  
Probability of voting in 2015 increases with age and better job conditions.

### **5.5 Information about the residency in Norway (x2)**

This is the second data set and it is referred to it as (x2) data set. It contains information about the respondent’s residency in Norway. x2 has six variables; originally from, the respondent’ status in Norway, the respondent’s residency period in Norway, the respondent’s residency status in Norway, the respondent’s address, the respondent’s Norwegian language level.

#### **5.5.1 Respondent’s residency variables general description**

- Originally from; the “originally from” variable is classified to 7 levels, as follows
  - ❖ Level 1: Europe
  - ❖ Level 2: Asia
  - ❖ Level 3: Middle East
  - ❖ Level 4: Africa
  - ❖ Level 5: North America
  - ❖ Level 6: Latin America
  - ❖ Level 7: Australia

Figure (5.9) shows The “participant’s home country” variable distribution. It distributes from level (1) to level (7) . The majority of the participants come from the Middle East.

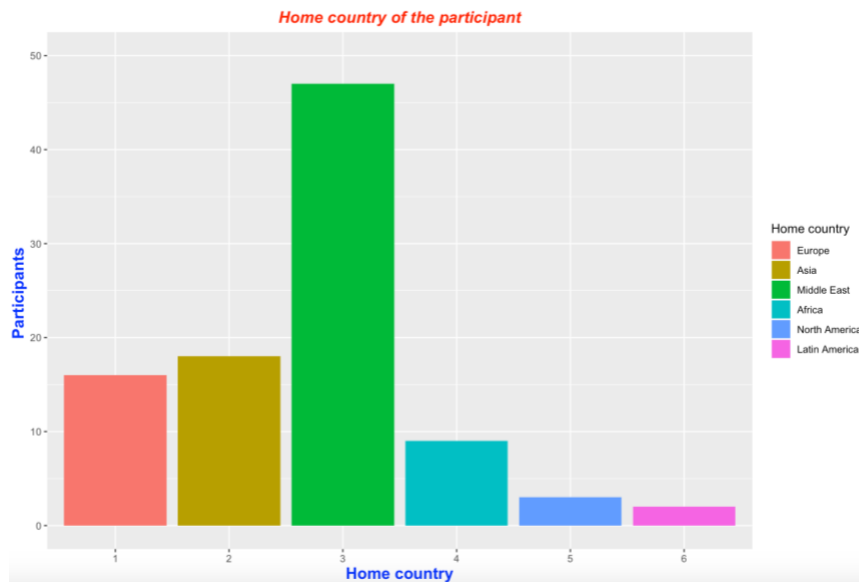


Figure (5.9): The “participant’s home country” variable distribution among participants

- **Status**

The “Status” variable is classified to two levels, as follows:

- ❖ Level 1: I was born in Norway, both parents with immigrant background.
- ❖ Level 2: I moved to Norway.

Figure (5.10) shows the “participant’s status in Norway” variable distribution. It distributes from level 1 to level 2. The majority of the participants moved to Norway, they weren’t born in Norway.

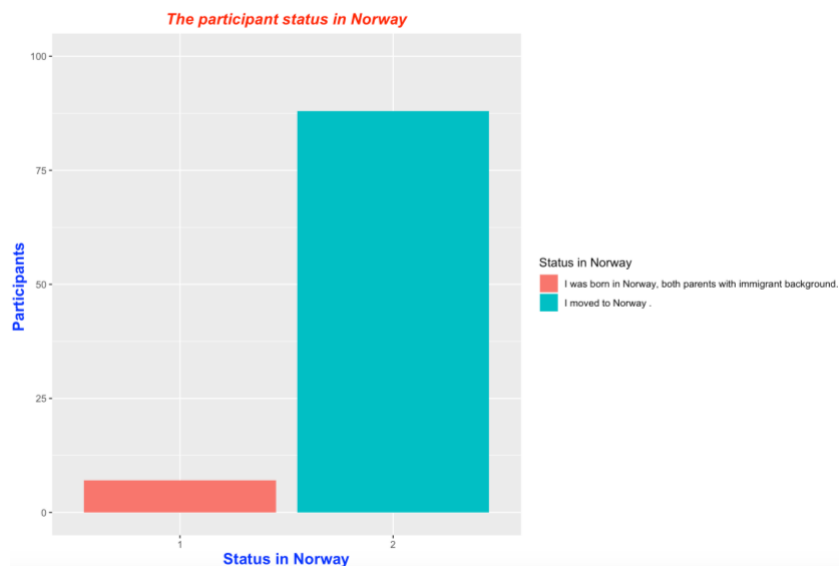


Figure (5.10): The “participant’s status in Norway” variable distribution among participants

- Residency period; the “Residency period “variable is classified from level (1) to (7) levels, as follows:

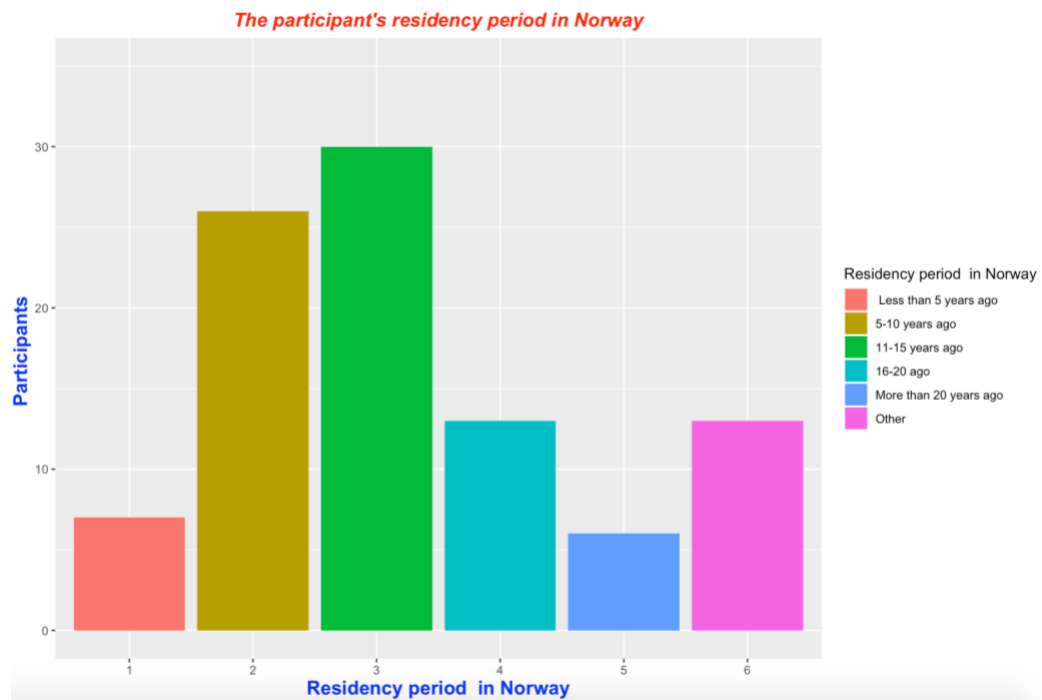


Figure (5.11): The “Residency period” variable distribution among participants

- ❖ Level 1: Less than 5 years ago
- ❖ Level 2: 5-10 years ago
- ❖ Level 3: 11-15 years ago
- ❖ Level 4: 16-20 ago
- ❖ Level 5: More than 20 years ago
- ❖ Level 6: Other

Figure (5.11) shows the “participant’s residency period” variable distribution. It distributes from level 1 to level 6. The majority of the participants have residency period in Level 3.

- Residency permit type

The “Residency permit type “variable is classified (1) to (9) levels, as follows:

- ❖ Level 1: Norwegian citizen
- ❖ Level 2: Work/entrepreneur
- ❖ Level 3: Study
- ❖ Level 4: Family reunification
- ❖ Level 5: Permission for family members of EU/EEA nationals
- ❖ Level 6: Permanent resident
- ❖ Level 7: Asylum-seeker
- ❖ Level 8: Refugee
- ❖ Level 9: Other

Figure (5.12) shows the “participant’s residency period” variable distribution. It distributes from level 1 to level 9. The majority of the participants have residency period in level 1

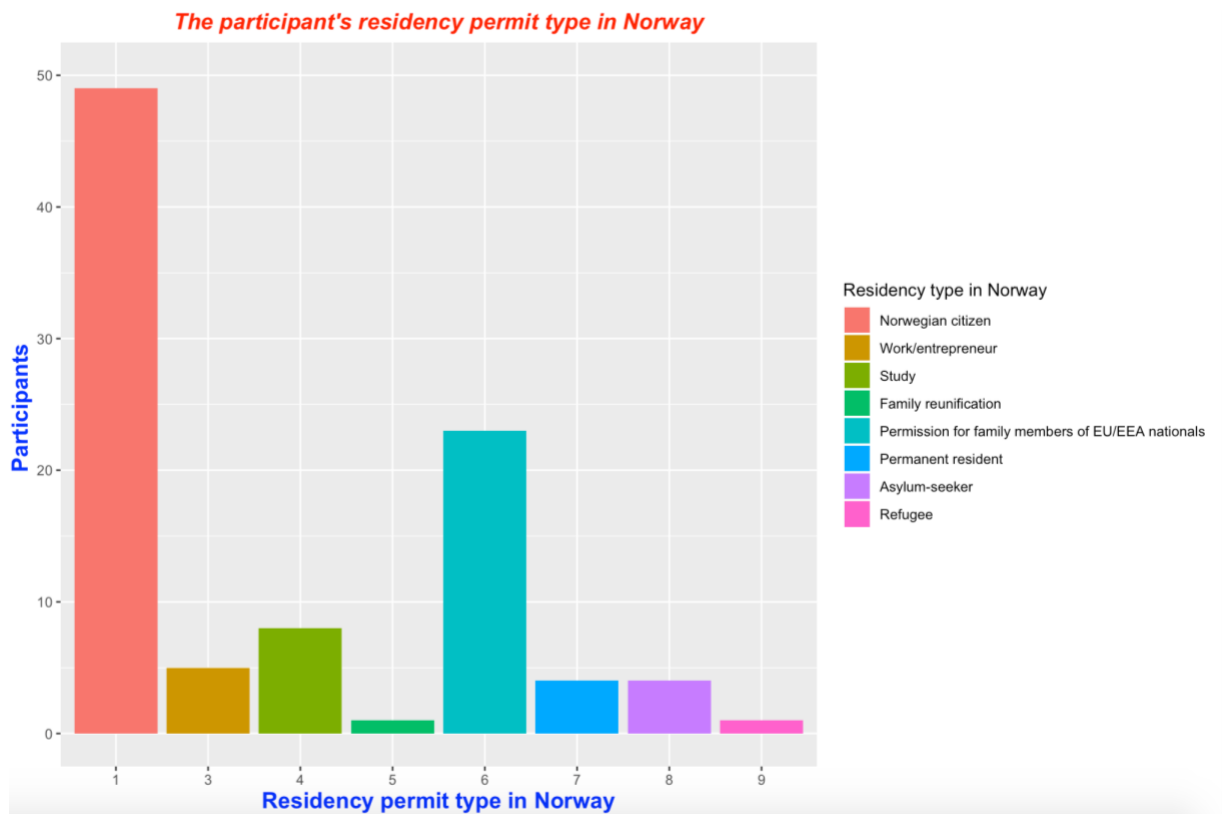
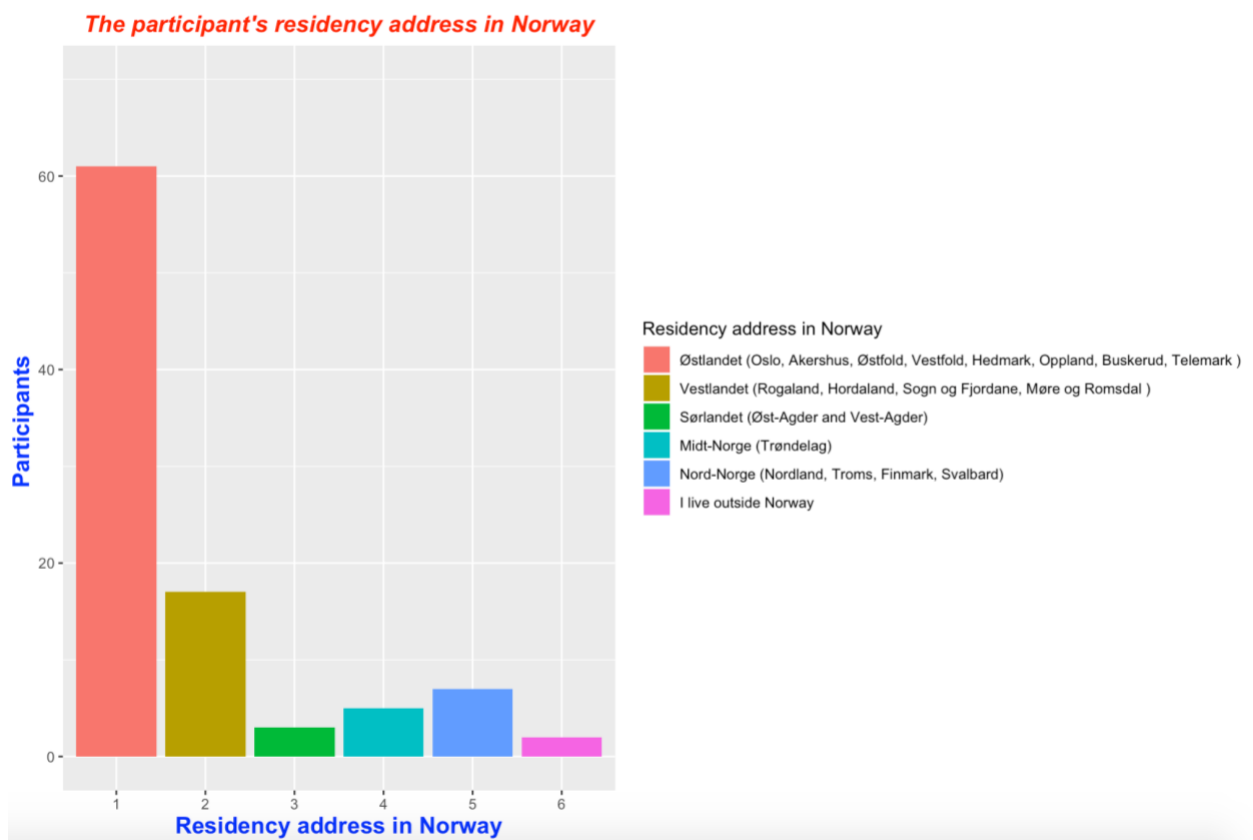


Figure (5.12): The ”Residency permit type in Norway” variable distribution among participants

- I live in; the “I live in “variable is classified to 6 levels, as follows:
  - ❖ Level 1: Østlandet (Oslo, Akershus, Østfold, Vestfold, Hedmark, Oppland, Buskerud, Telemark )
  - ❖ Level 2: Vestlandet (Rogaland, Hordaland, Sogn og Fjordane, Møre og Romsdal)
  - ❖ Level 3: Sørlandet (Øst-Agder and Vest-Agder)
  - ❖ Level 4: Midt-Norge (Trøndelag)
  - ❖ Level 5: Nord-Norge (Nordland, Troms, Finmark, Svalbard)
  - ❖ Level 6: I live outside Norway



Figure(5.13):The “Residency address in Norway” variable distribution among participants

Figure (5.13) shows the “I live in” variable distribution. It distributes from Level (1) to Level (6). The majority of the participants have “ I live “ variable in Level (1).

- Norwegian language level; the “Norwegian language level” variable is classified to 8 levels, as follows:
  - ❖ Level 1: Proficient” First language”, C2
  - ❖ Level 2: Advanced, C1
  - ❖ Level 3: Upper-intermediate, B2
  - ❖ Level 4: Intermediate, B1
  - ❖ Level 5: Pre-intermediate, A2
  - ❖ Level 6: Elementary, A1
  - ❖ Level 7: Beginner, A0-A1
  - ❖ Level 8: None

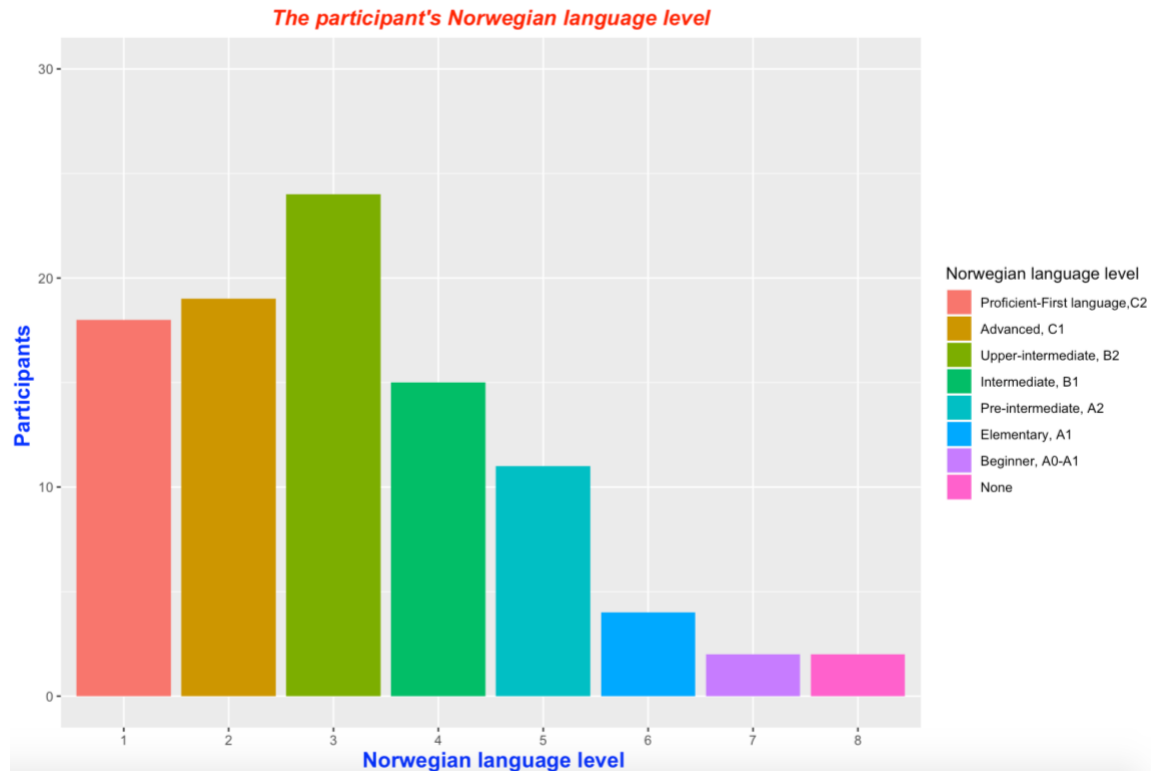


Figure (5.14 ):The “Norwegian language level” variable distribution among participants

Figure(5.14) shows “Norwegian language level” variable distribution. It distributes from level (1) to level (8). The majority of the participants have “Norwegian language level” variable in level 3.



### **5.5.2 Correlation**

By finding the correlation relation between the respondent's residency variables, the important results are :

- ❖ Type of residency permit correlates negatively with the residency period
- ❖ Norwegian language level increases with stronger residency permit type and longer residency period

### **5.5.3 Collinearity and multicollinearity**

Here, the calculated ration between the maximum and minimum eigenvalues = 6.51. This indicates that the multicollinearity is not significant issue between the x2 variables.

### **5.5.4 Correlation relation between residency variables and voting in 2017 national elections variables**

The highest positive correlation is with residency period = 0.53,.The highest negative correlation is with type of residency permit = - 0.53

Voting in 2017 increases for longer residency period and stronger residency permit type

### **5.5.5 Correlation relation between residency information variables and voting in 2015 local elections variables**

The highest positive correlation is with residency period= 0.49. The highest negative correlation is with Norwegian language level= - 0.45

Voting in 2015 increases for longer residency period and higher Norwegian level.

## 5.6 The Political Participation dataset variables (x3)

This is the third data set, it is referred to it as (x3) data set. It is the core of the questionnaire and contains information about the political participation in general. This data set contains nine explanatory variables; why did she vote, why didn't she vote? If she used to vote at the original country? the respondent's political participation, If she has ever been a candidate or member of political party, if she has ever been a member or an interest group that working for a change socially or politically. In the last 12 months, has she volunteered her time for a group or organization other than a political party? in the last 12 months, has she taken a part in a demonstration? in the last 12 months, has she signed a petition?

The first three questions appears to specific respondents based on their answers to some previous questions( it is described at the beginning of this chapter)

### 5.6.1 Respondent's political participation variables general description

- Political participation; is classified from level (1) to level (5), as follows:
  - ❖ Level 1: Very interested
  - ❖ Level 2: Fairly interested
  - ❖ Level 3: Only Vaguely interested

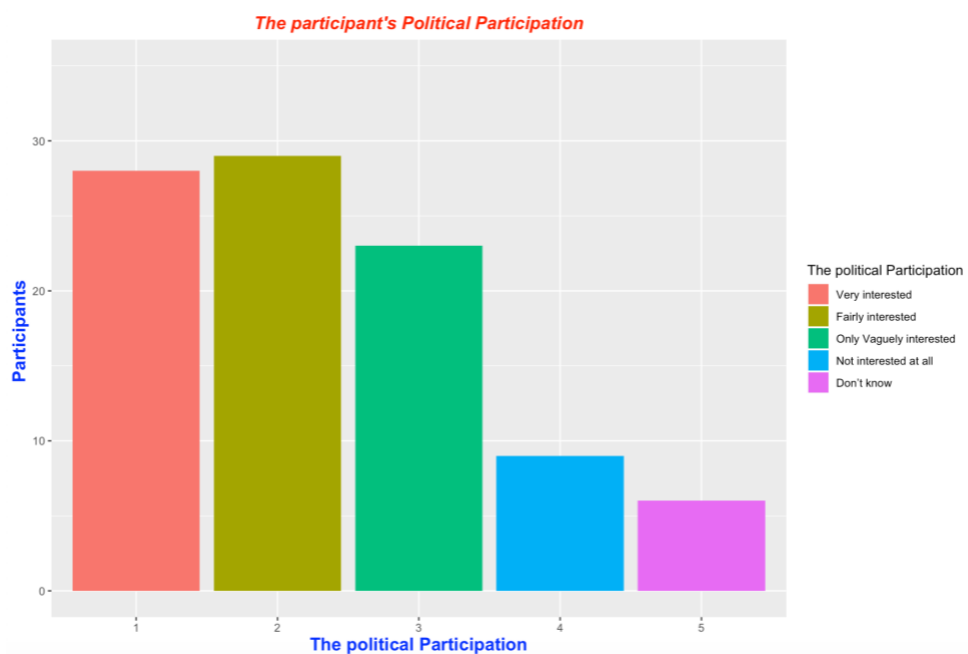


Figure (5.15): The “Political participation” variable distribution among participants

- ❖ Level 4: Not interested at all
- ❖ Level 5: Don't know

Figure (5.15) shows the “Political participation” variable distribution. It distributes from level 1 to level 5,. The majority of the participants have “Political participation “variable in level (2).

- Why did the respondent vote; is classified to 7 levels, as follows:
  - ❖ Level 1: I wanted to support a specific party or candidate
  - ❖ Level 2: I wanted to vote against a specific party or candidate
  - ❖ Level 3: I was concerned about a particular issue or issues raised I think voting is important
  - ❖ Level 4: I wanted to have my voice heard
  - ❖ Level 5: I always vote when I am eligible to vote



Figure (5.16): The “Why did the respondent vote?” variable distribution among participants

- ❖ Level 6: Family or friends persuaded me to vote
- ❖ Level 7: Other

Figure (5.16 ) shows the modified “Why did the respondent vote?” variable distribution. It distributes from level 1 to level 7. The majority of the participants have chosen the “why did vote” variable level(1).

- Why didn’t the participant vote?; is classified to 14 levels, as follows:
  - ❖ Level 1: Not eligible
  - ❖ Level 2: I didn't know where or when to vote
  - ❖ Level 3: I didn't think I was on the voters' list.
  - ❖ Level 4: I didn't think my vote would matter to the outcome.
  - ❖ Level 5: I didn't know enough about the candidates, the parties, or the issues.
  - ❖ Level 6: I didn't like the candidates or parties.
  - ❖ Level 7: I wasn't concerned about the issues that were raised.
  - ❖ Level 8: Personal circumstances on election day
  - ❖ Level 9: I didn't have enough time on election day.
  - ❖ Level 10: I just wasn't interested.
  - ❖ Level 11: I was scared to vote.
  - ❖ Level 12: I didn't speak Norwegian well enough.
  - ❖ Level 13: I don't think any politicians can be trusted.
  - ❖ Level 14: Other

Figure (5.17) shows “Why didn’t the participant vote? “variable distribution. It distributes from level 1 to level 14. The majority of the participants have the modified “Why didn’t the participant vote? “variable in Level (1) and Level (5).

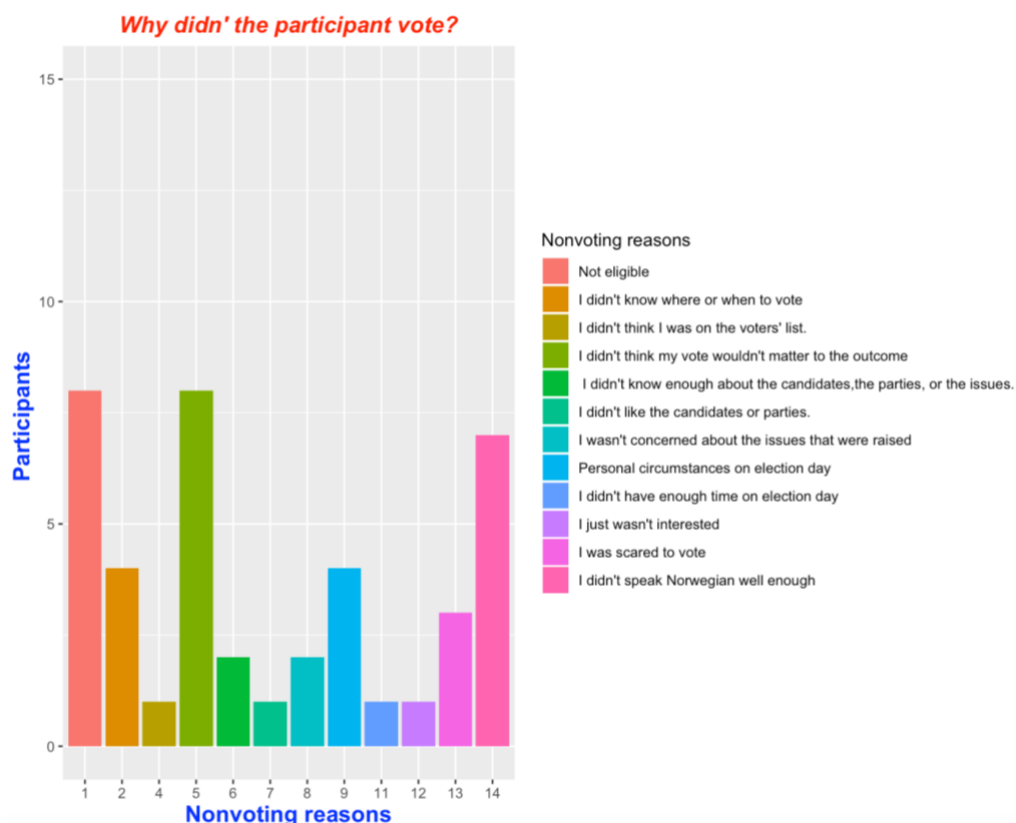


Figure (5.17): The “Why didn’t the participant vote?” variable distribution among participants

- Voting at the elections in home country; this variable is classified to 6 levels, as follows:
  - ❖ Level 1: There were elections and I usually voted
  - ❖ Level 2: There were elections and I sometimes voted
  - ❖ Level 3: There were elections, but I didn't vote
  - ❖ Level 4: There were elections, but I was not eligible to vote
  - ❖ Level 5: There were no elections at all in my country of origin
  - ❖ Level 6: There weren't regular elections

Figure (5.18) shows “Voting at the election in home country” variable distribution. It distributes from level 1 to level 6. The majority of the participants have the modified “Why didn’t the participant vote?” variable in Level (1).

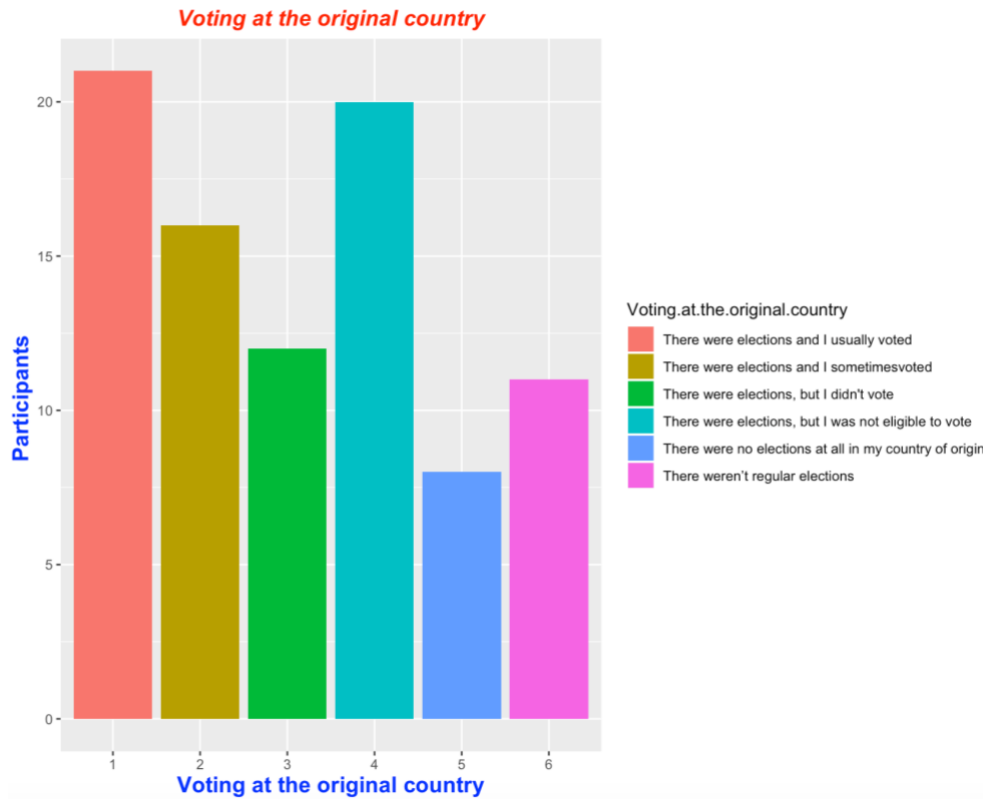
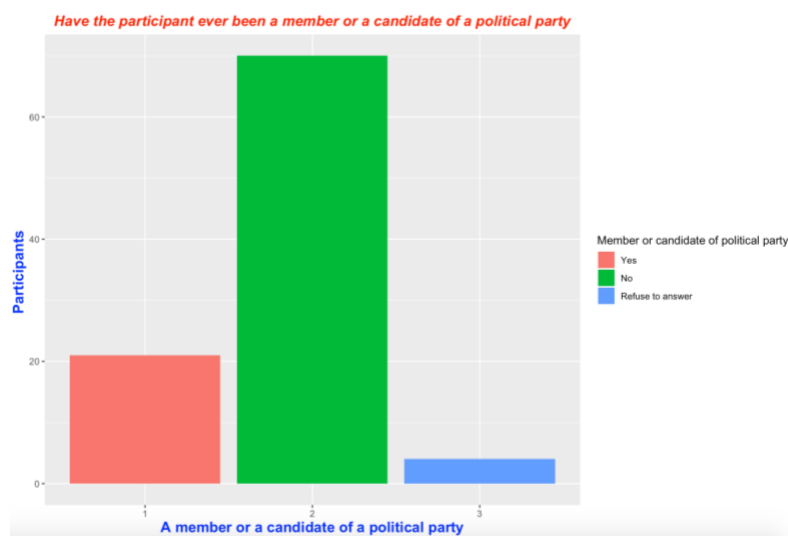


Figure (5.18): The “Voting at the election in home” variable distribution among participants

- A member or a candidate of a political party; this variable is classified to 3 levels;



Figure(5.19): “A member or a candidate of a political party” variable distribution

- ❖ Level 1: Yes
- ❖ Level 2: No
- ❖ Level 3: Refuse to answer

Figure (5.19) shows “A member or a candidate of a political party” variable distribution. It distributes from level 1 to level 3. The majority of the participants have “A member or a candidate of a political party “ variable in level (2).

- Member of an interest group working for social or political change variable; is classified to 3 levels, as follows:
  - ❖ Level 1: Yes
  - ❖ Level 2: No
  - ❖ Level 3: Refuse to answer

Figure (5.20) shows “A member of an interest group working for social or political change” variable distribution. It distributes from level 1 to level 3. The majority of the participants have “A member of an interest group working for social or political change “variable in level (2).

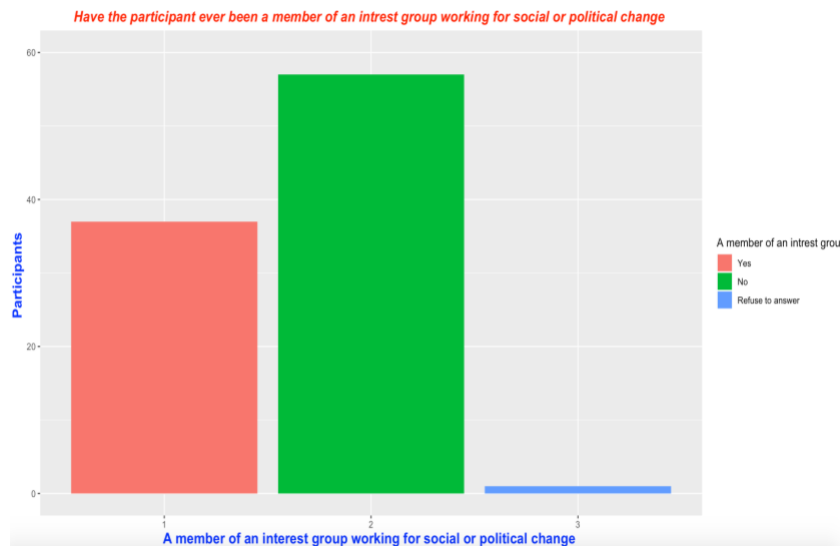


Figure (5.20): “A member of an interest group working for social or political change” variable distribution among participants

- A volunteer for a group or organization other than a political party, in the last 12 months; is classified to 3 levels, as follows: Level
  - ❖ Level 1: Yes
  - ❖ Level 2: No
  - ❖ Level 3: Refuse to answer



Figure (5.21): “A volunteer for a group or organization other than a political party” variable distribution among participants

Figure (5.21) shows “A volunteer for a group or organization other than a political party, in the last 12 months” variable distribution. It distributes from level 1 to level 3. The majority of the participants have “A volunteer for a group or organization other than a political party, in the last 12 months” variable in Level 2.



- Taking a part in a demonstration, in the last 12 months; is classified to 3 levels, as follows:
  - ❖ Level 1: Yes
  - ❖ Level 2: No
  - ❖ Level 3: Refuse to answer

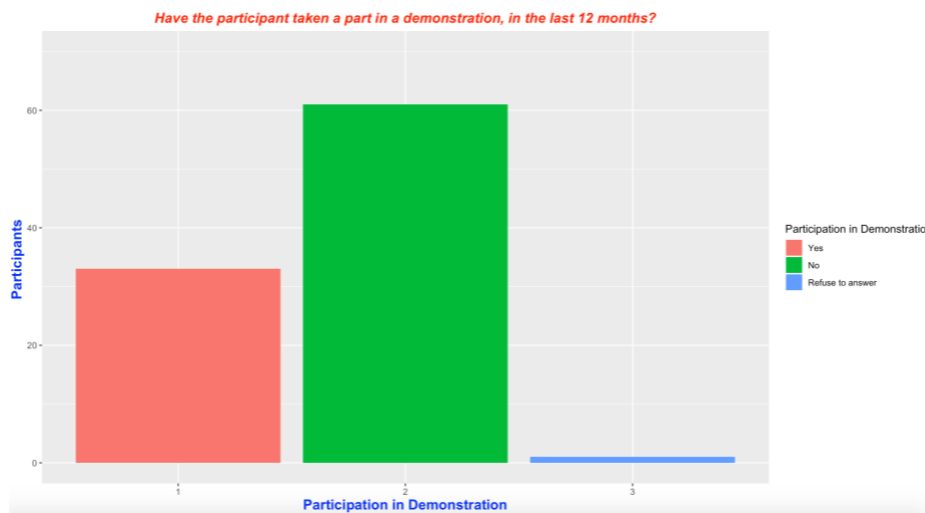


Figure (5.22) : “Taking a part in a demonstration, in the last 12 months” variable distribution

Figure (5.22) shows “Taking a part in a demonstration, in the last 12 months” variable distribution. It distributes from level 1 to level 3. The majority of the participants have “Taking a part in a demonstration, in the last 12 months “variable in Level 2.

- Signing a petition, in the last 12 months’ is classified to 3 levels, as follows:
  - ❖ Level 1: Yes
  - ❖ Level 2: No
  - ❖ Level 3: Refuse to answer

Figure (5.23) shows “Signing a petition, in the last 12 months” variable distribution. It distributes from level 1 to level 3. The majority of the participants have “Signing a petition, in the last 12 months “variable in Level (2).

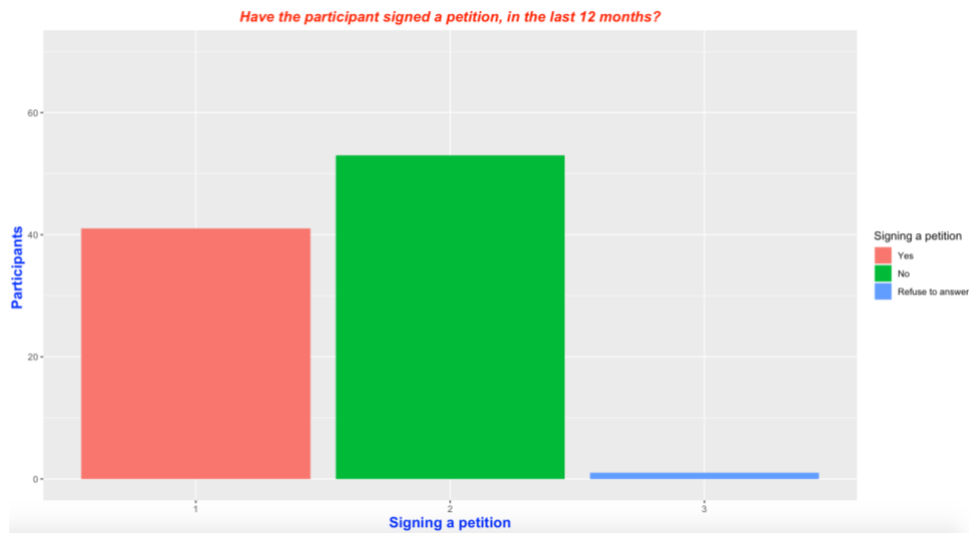


Figure (5.23): “Signing a petition, in the last 12 months” variable distribution among participants

It is important to mention that based on the questionnaire structure, the dependent variables “voting in 2017 national elections” and “voting in 2015 local elections” belong to the third section in the questionnaire. These were excluded from x3 at the data analysis. But it can be described here as a part of the data explanatory analysis to have impression about its values.

- Voting in the 2015 local elections: (Dependent variable and excluded from x3 data set variables); is classified to 2 levels, as follows:
  - ❖ Level 1: No
  - ❖ Level 2: Yes

Figure (5.24) shows the “Voting in the 2015 local elections” variable distribution. The majority of the participants have the “Voting in the 2015 local elections” variable in Level 1.

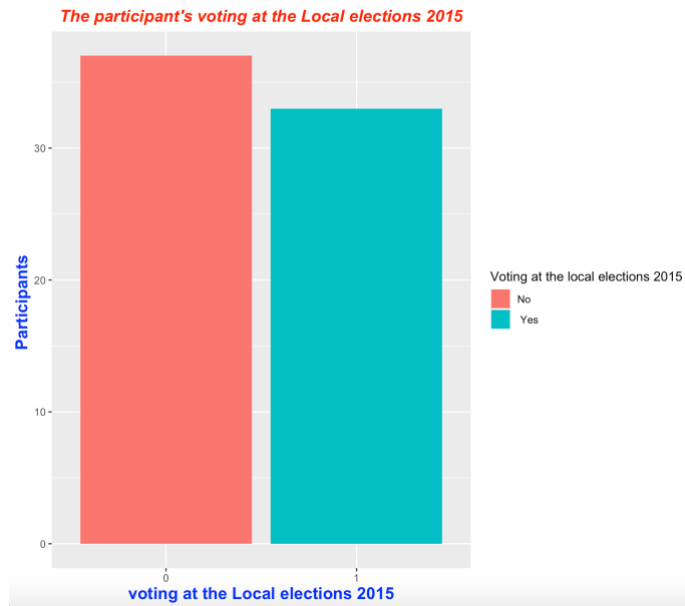
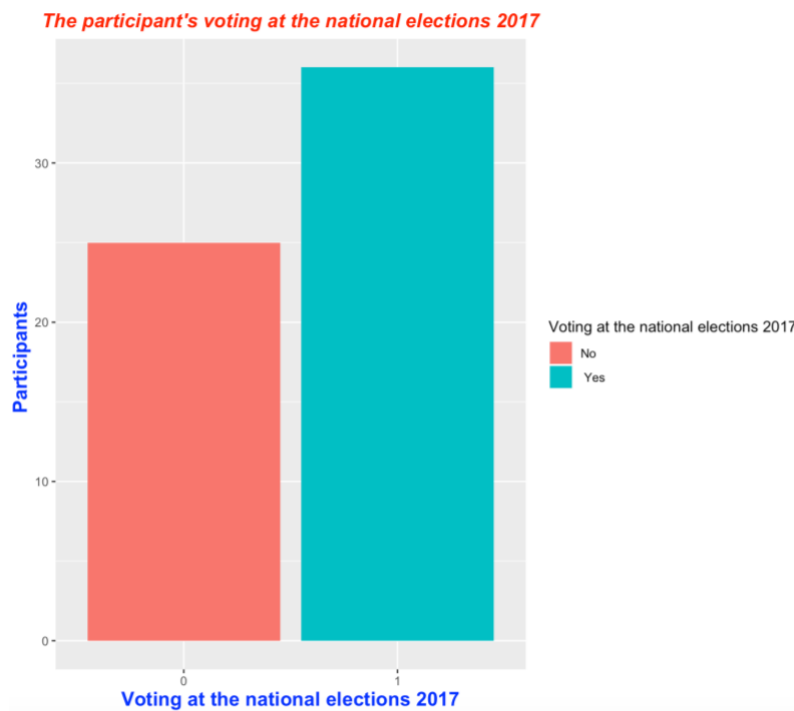


Figure (5.24): "Voting in the 2015 local elections" variable distribution among participants

- Voting in the 2017 national elections (Dependent variable and excluded from x3 data set variables); is classified to 2 levels, as follows:
  - ❖ Level 1: No
  - ❖ Level 2: Yes



Figure(5.25): "Voting in the 2017 national elections" variable distribution among participants

Figure (5.25) shows the “Voting in the 2017 national elections” variable distribution. The majority of the participants have the “Voting in the 2017 national elections” variable in level 2.

### **5.6.2 Correlation**

By finding the correlation relation for x3 data set variables, some the following results are found:

Voting at the original country: has the positive correlation with “interested at the political participation” and negative correlation with “In the last 12 months, has volunteered her time for a group or organization other than a political party”. Interested in political participation: has negative correlation with “why did vote? Why didn’t vote: has negative correlation with “Why did vote?”.

Generally this indicates, the stronger reasons to why did vote, the weaker reasons to why not vote, the more interested with politics and low participation in election at home country.

### **5.6.3 Collinearity and multicollinearity**

Here, the calculated ration between the maximum and minimum eigenvalues = 18.67. This indicates that the multicollinearity is not significant issue between the x3 variables, since the value is less than 100

### **5.6.4 Correlation relation between political participation variables and “voting in 2017 national elections” variable**

The highest positive correlation is with “Why did vote”. Voting in 2017 increase with the stronger reasons for why did vote.

### **5.6.5 Correlation relation between political participation variables and “voting in 2015 local elections” variable**

The highest positive correlation is with “Why did vote”. The highest negative correlation is with “Why didn’t vote”. Voting in 2015 increase with the stronger reasons for why did vote, and weaker reasons to why didn’t vote.

## 5.7 The surrounded environment and the political view (x4)

This is the fourth data set, it is referred to it as (x4) data set. It contains four variables about the effect of the surrounded environment on the participant political view. The variables are; How often did the family used to discuss politics, when the respondent was a child? How often do you discuss politics now?, how many immigrant friends does the respondent have? and the degree that family, friends and relative affect the respondent's choice of the political party.

### 5.7.1 The surrounded environment and the political view variables general description

- How often did your family used to discuss politics, when you were a child?; is classified to 5 levels, as follows:
  - ❖ Level 1: Frequently
  - ❖ Level 2: Occasionally
  - ❖ Level 3: Seldom
  - ❖ Level 4: Never
  - ❖ Level 5: Don't know

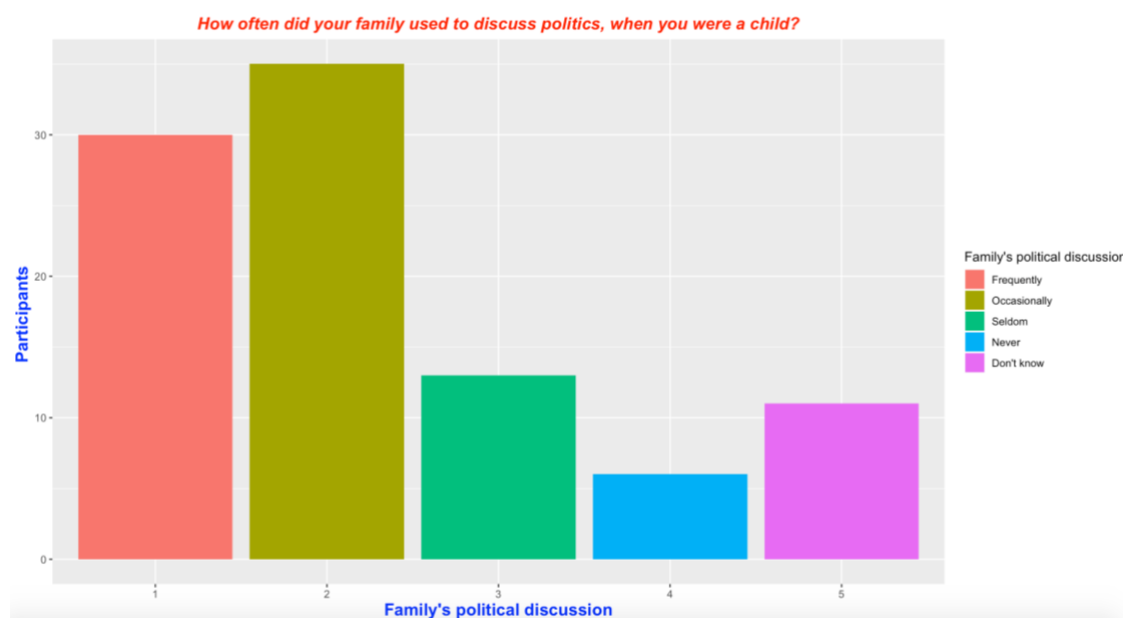


Figure (5.26): “How often did your family used to discuss politics” variable distribution among participants

Figure (5.26) shows “How often did your family used to discuss politics” variable distribution. It distributes from level 1 to level 5. The majority of the participants have “How often did your family used to discuss politics “variable in level (2).

- How often do you discuss politics now?; is classified to 4 levels, as follows:
  - ❖ Level 1: Every day
  - ❖ Level 2: A few times a week
  - ❖ Level 3: Never
  - ❖ Level 4: Don’t know

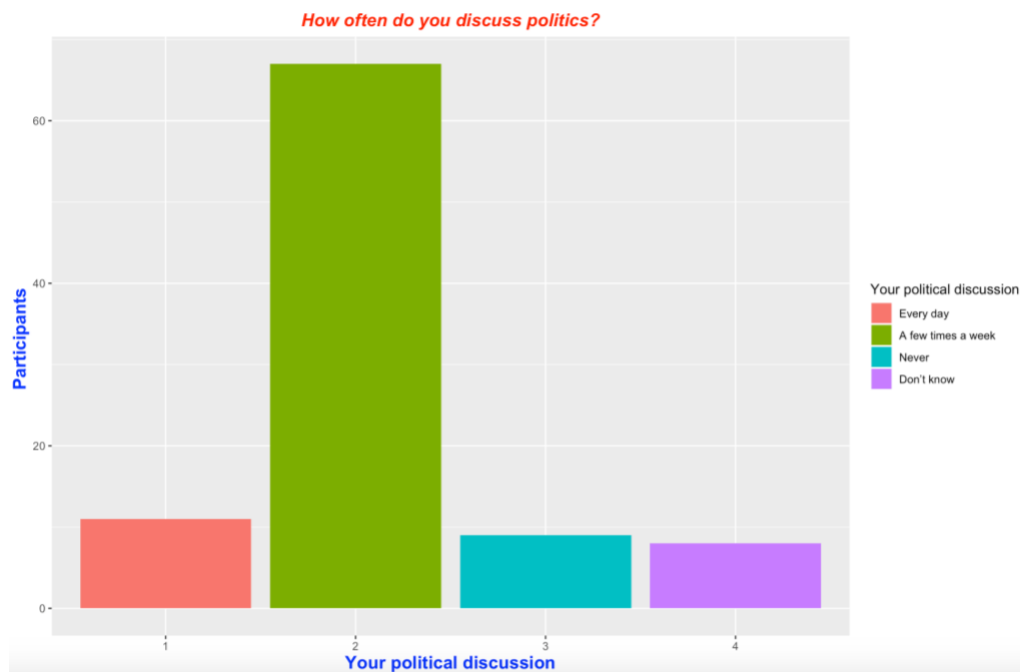


Figure 5.27: “How often do you discuss politics now?” variable distribution among participants

Figure 5.27 shows “How often do you discuss politics now?” variable distribution. It distributes from level 1 to level 4. The majority of the participants have “How often do you discuss politics? “variable in level (2).

- How many of your friends have an immigrant background? is classified to 7 levels, as follows :

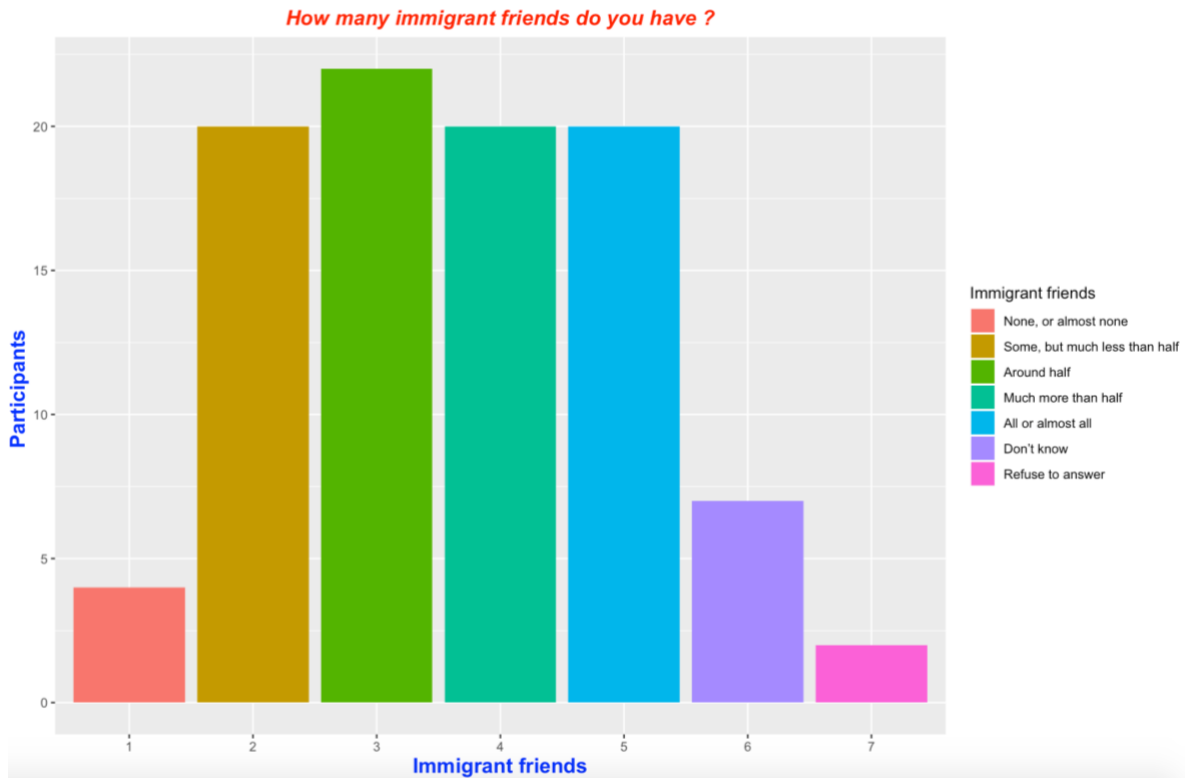


Figure 5.28: “How many of your friends have an immigrant background?” variable distribution

- ❖ Level 1: None, or almost none
- ❖ Level 2: Some, but much less than half
- ❖ Level 3: Around half
- ❖ Level 4: Much more than half
- ❖ Level 5: All or almost all
- ❖ Level 6: Don't know
- ❖ Level 7: Refuse to answer

Figure (5.28) shows that the number of participant who have their “How many of your friends have an immigrant background?” answer in level 1 are: 7. The majority answered level (3)

- To what degree did you experience pressure from your closed people with your choice of a political party? is classified to 5 levels, as follows:

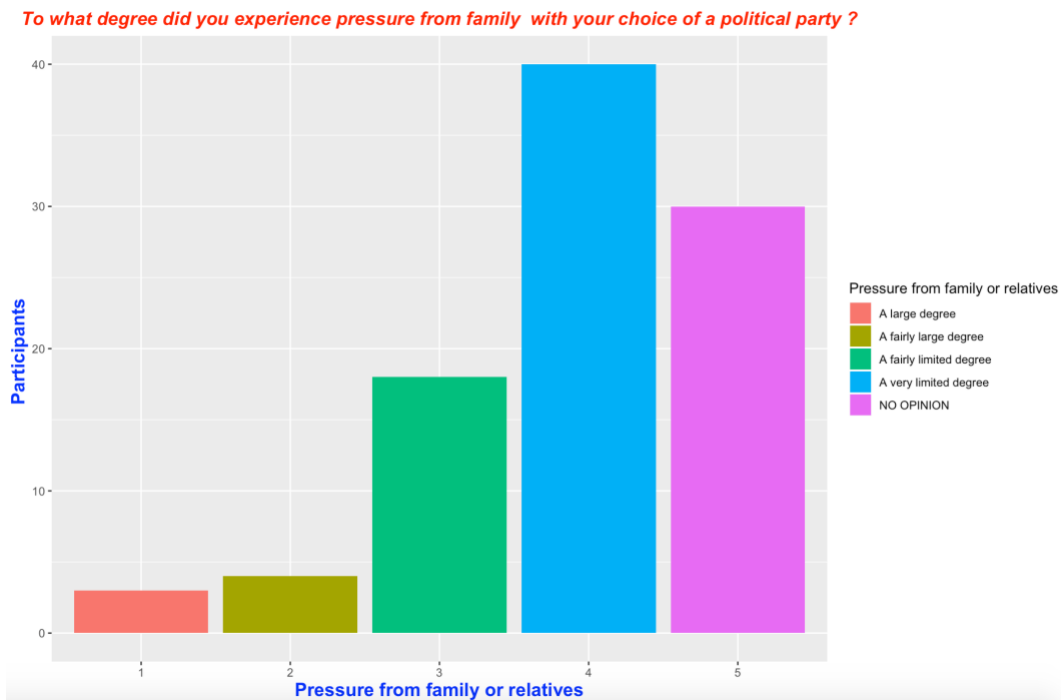


Figure (5.29): “pressure from your with your choice of a political party” variable distribution

- ❖ Level 1: A large degree
- ❖ Level 2: A fairly large degree
- ❖ Level 3: A fairly limited degree
- ❖ Level 4: A very limited degree
- ❖ Level 5: No opinion

Figure (5.29) shows that the variable distribution from level(1) to level (5). The majority answered level (4)



### **5.7.2 Correlation**

Family discussed politics increases with how often discuss politics now and less immigrant friends.

### **5.7.3 Collinearity and multilinear**

Here, the calculated ration between the maximum and minimum eigenvalues = 3.32. This indicates that the multicollinearity is not significant issue between the x4 variables, since the value is less than 100

### **5.7.4 Correlation relation between X4 variables and “voting in 2017 national elections variable”**

Here, voting in 2017 increases with more often family discussed politics.

### **5.7.5 Relation between X4 variables and “voting in 2015 local elections variable”**

Here, voting in 2015 increases with more often family discussed politics and less pressure on the opinion.

## **5.8 Political view data set (x5)**

The fifth and the last data set. It is referred to it as x5 data set. It contains variable about the general political view for the respondent's. It contains 22 variables. Starts with "e\_po\_15" variable to "e\_po\_36", see (Table 1. Appendix B).

All of the 22 questions are scaled questions. Having the following answers:

Strongly disagree, disagree, somewhat disagree, neither disagree nor agree, somewhat agree, agree, strongly agree. The answers start from strongly disagree=1 to strongly agree =7

### **5.8.1 Collinearity and multilinear**

Here, the calculated ration between the maximum and minimum eigenvalues = 27.19. This indicates that the multicollinearity is not significant issue between the x4 variables, since the value is less than 100.

# Chapter 6

## Data Modelling

### Summary

In this chapter the used statistical methods for modelling and inference are presented. The model selection methods are discussed as well. The chapter contains four sections; the first section discusses logistic regression modelling. The second section presents the principle components analysis. While the third section views the principle components logistic regression, moreover the data imputation and standardising will be presented. Furthermore the fourth section discusses model selections methods.

### 6.1 Logistic Regression

#### 6.1.1 linear regression

The statistical technique that is used to find the linear relation between a set of (x,y) pairs is called simple linear regression. It is used for fitting a straight line that explains this linear relation. x here is known as the explanatory variable while y is known as the response. The intercept and the slope of the fitted straight line are selected to minimize the sum of squared differences between the measured and fitted response's values. <sup>77</sup>

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad (6.1)$$

where; y and epsilon are random variables in this model. beta-s parameters and x explanatory variable that is assumed to be known. In the simple linear regression there are two important structures; the mean and the error where ;

---

<sup>77</sup> Cook, D et al. 2001. "Binary Response and Logistic Regression Analysis." In *Beyond Traditional Statistical Methods*

The mean structure:

$$E(y_i | x) = \beta_0 + \beta_1 x_i \quad (6.2)$$

And ; the error structure:

$$\varepsilon_i \sim N(0, \sigma^2) \quad (6.3)$$

It is worth mentioning that the explanatory variable that is used to predict the response is numerical measurement and the response as well . When there is a set of explanatory variables or predictors, for example instead of  $x$ , there is a denoted vector  $(x_1, x_2, x_3, x_4, \dots, x_k)$ , the statistical technique is used to fit the linear relation between it and the model is known as multiple regression.

### 6.1.2 Logistic regression (LR)

When the response  $y$  has another data structure, dichotomous (binary), there is another type of regression used to explain the relation between the response and the explanatory variables. It is called logistic regression(LR).<sup>78</sup> Logistic regression is used a predictive analysis as well. It is used to predict set of response's characteristics that depend on the occurrence or non-occurrence of the event of interest. In addition logistic regression hasn't the assumption that are adopted by other regression types. In the case that the response has binary values; for example absence or a presence of a condition or occurrence or not of an event of interest ;  $y_i = 1$  ( as a success trial ) and  $y_i = 0$  ( as a failure trial). It is known as a dichotomous response variable, and can be re-expressed as<sup>79</sup>:

$$\text{Prob}(y_i = 1) = \pi_i \quad (6.4)$$

$$\text{Prob}(y_i = 0) = 1 - \pi_i \quad (6.5)$$

With the binary response , the expected value is;

---

<sup>78</sup> Statistics solutions. "What Is Logistic Regression?"

<sup>79</sup> Cook, D et al. 2001. "Binary Response and Logistic Regression Analysis." In Beyond Traditional Statistical Methods,.

$$E(y_i) = 0 \cdot (1 - \pi_i) + 1 \cdot \pi_i = \pi_i \quad (6.6)$$

Denoting  $x$  vector  $(x_1, x_2, x_3, x_4, \dots, x_k)$  as the predictor or the explanatory variable vector, the binary logistic regression is here <sup>80</sup>;

$$\text{Prob}\{y = 1|x\} = [1 + \exp(-x\beta)]^{-1}. \quad (6.7)$$

Where :

$$x\beta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (6.8)$$

$\beta$  is the regression parameter vector which is estimated by using maximum likelihood method.  $\beta_0, \beta_1 x_1, \beta_2 x_2, \dots, \beta_k x_k$  are the regression coefficients as well .

Since equation(6.7) is a ratio equation, can be written this way ;

$$\text{Prob}\{y = 1|x\} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)}} \quad (6.9)$$

By taking the log to the ration of  $P(y=1)$  to  $(1-P(y=1))$  as in equation (6.10)

$$g(x) = \log \frac{P(y=1)}{(1-P(y=1))} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (6.10)$$

A linear model in  $x_i$  is resulted, which shows that the logistic regression model has a hidden linear regression model.

### 6.1.3 Logistic regression in R

The logistic regression is implemented in the open resource software program R using the function `glm()`; as follows<sup>81</sup> ;

$$\text{Glm}(\text{formula}, \text{family}, \text{na.action}, \dots)$$

The formula contains here the response variables and the explanatory variables. The family can be binomial, gaussian or any other distribution. The default link for the binomial family is `logit`

---

<sup>80</sup> Harrel Jr., Frank E. 2015. "Regression Modeling Strategies - With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis." R Software.

<sup>81</sup> Manning, Christopher. 2007. "Logistic Regression ( with R )." *Changes*.

## 6.2 Principle component analysis (PCA)

Principle component technique (PCA) is used for data dimensional reduction while covering the maximum amount of data variability. It contributes in the identification of the principle direction of the data covariance as well. Furthermore, this technique can PCA demands creating new set of variables that are called the principle components. Each component is a linear combination of the original data variables.

From the computational side, PCA can be calculated by finding the eigenvectors and eigenvalues of the data covariance matrix. Taking into consideration that the eigenvector with highest eigenvalue is the direction of the greatest data covariance, the one with the next highest eigenvalue is the direction of the next greatest data covariance and so on. The computation of eigenvectors and eigenvalues can be explained as follows(Gillies, Duncan Fyfe. 2006):<sup>82</sup>

Define  $\mathbf{A}$  as  $n \times n$  matrix.  $\mathbf{I}$  is  $n \times n$  identity matrix;

$$\text{determinant}(\mathbf{A} - \lambda\mathbf{I}) = |(\mathbf{A} - \lambda\mathbf{I})| = 0 \quad (6.11)$$

The roots of equation (6.11) define the eigenvalues of  $\mathbf{A}$ . This equation has  $n$  roots and called the characteristic equation or characteristic polynomial. By letting  $\lambda$  be the eigenvalue of  $\mathbf{A}$ . By denoting  $\mathbf{x}$  as a vector ;

$$\mathbf{Ax} = \lambda\mathbf{x} \quad (6.12)$$

So  $\mathbf{x}$  is the eigenvector of  $\mathbf{A}$  that is associated with the eigenvalue  $\lambda$ .  $\mathbf{x}$  has no unique solution since it is defined here as a direction vector that can be scaled by any magnitude. Furthermore for giving  $\mathbf{x}$  numerical solution and based on equation (6.11), one of its elements needs to have an arbitrary value, for example to value 1. So for solving the other elements, this can contribute in giving simultaneous equations. Normalizing the final values leads to having  $\mathbf{x}$  with length one, that donate  $\mathbf{x} \cdot \mathbf{x}^T = 1$ .

---

<sup>82</sup> Gillies, Duncan Fyfe. 2006. "Lecture 15: Principal Component Analysis." *Intelligent Data Analysis*.

By supposing that  $A$  is a  $3 \times 3$  matrix, that has eigenvectors  $x_1, x_2, x_3$  and eigenvalues  $\lambda_1, \lambda_2, \lambda_3$ . And referring to equation (6.11); this follows to :

$$Ax_1 = \lambda_1 x_1, \quad Ax_2 = \lambda_2 x_2, \quad Ax_3 = \lambda_3 x_3 \quad (6.13)$$

And to;

$$A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \quad (6.14)$$

By defining :

$$\Phi = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{and,} \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

This leads to

$$A \Phi = \Phi \Lambda \quad (6.15)$$

As mentioning above, by normalizing the eigenvalues that are orthogonal to the unit magnitude, this follows to ;

$$\Phi \Phi^T = \Phi^T \Phi = I$$

That means ;

$$\Phi^T A \Phi = \Lambda \quad (6.16)$$

And;

$$A = \Phi \Lambda \Phi^T \quad (6.17)$$

This can be applied to PCA by defining matrix of eigenvectors  $\Phi$  as the linear transformation. The linear transformation given by  $\Phi$  transforms the data original variables to components that are uncorrelated. Moreover the new coordinate system has  $\Lambda$  as the data correlation matrix that has zeros for diagonal elements.

There are two key points should be mentioned and taken into account before applying the PCA method :

Data standardization: The principle components analysis is defined as the method that is used to reduce data dimensions by generating components that are linear combination of the original variables. Furthermore these components contains the majority of the data variability. But the principle components could show some of the undesired features and this happens when the original data variables has different measurement units. And this is because the generating of the components based on the covariance matrix which depends as well on the measurements unit. In case of the changing of the measurements units to any variables this results in changing at the covariance matrix so in the generated components as well. The best solution here is standardizing the data<sup>83</sup>. After standardising the data, the standardized data covariance matrix is merely defined as the correlation matrix of the original data set

Data imputation<sup>84</sup> : When the data has missing values, the data imputation is a strategy to fill the gaps before applying the principle components analysis. This strategy has two methods to fill the missing values that are; the easiest one is mean imputation method that is achieved by using the mean of the rest of the data values and fill in the missing values with the mean.

### **6.3 Principle component logistic regression (PCLR)**

The logistic regression is used to find the relation between a binary dependent variable and one or more of the explanatory variables. When the data set contains a large number of the explanatory variables, a statistical methods for reduce the data dimensions are used. The principle components analysis is one of the best methods for dimensional reduction. Firstly the PCA is applied to the original data set to compute the principle components and then the resulted components

---

<sup>83</sup> Jolliffe, Ian T., and Jorge Cadima. 2016. "Principal Component Analysis: A Review and Recent Developments." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*.

<sup>84</sup> Dray, Stéphane, and Julie Josse. 2015. "Principal Component Analysis with Missing Values: A Comparative Survey of Methods." *Plant Ecology*.

are used in fitting a logistic regression model. Using PCLR fitting model method contributes as well in improving the parameters estimation.<sup>85</sup>

## 6.4 Stepwise model selection

Fitting a model that describes data precisely with the presence of numerical stability and variables minimization, is a the best model that the majority looking for. Model selection techniques are used for avoiding the overfitting problem.

The criteria of variables' selection varies based on the purpose and the problem this model tries to explain. There are different kinds of the data selection procedures and methods such like; forward selection, backward elimination and stepwise selection

The stepwise model selection method is a mix of both forward selection method and backward elimination method. It starts same as forward selection by examining the significance of the effect at the model, but adding it doesn't guarantee that it will remain. Effects added to the model then examined all again and backward elimination is implemented to remove those that don't met the significance level of staying. Each forward step is followed by backward step at least. The process continues until having the best model.<sup>86</sup>

---

<sup>85</sup> Aguilera, Ana M., Manuel Escabias, and Mariano J. Valderrama. 2006. "Using Principal Components for Estimating Logistic Regression with High-Dimensional Multicollinear Data." *Computational Statistics and Data Analysis*.

<sup>86</sup> Zhang, heng. 2016. "Variable Selection with Stepwise and Best Subset Approaches." *Annals of Translational Medicine*.



# Chapter 7

## Data analysis results

### Summary

This chapter summarizes the results of analysing the data that was introduced in Chapter 5. The used methods for the data analysis here were presented in Chapter 6. The first section shows the analysis results for fitting binomial multiple logistic regression model. While the second section presents the principle component analysis results. The third section summarizes the results of fitting principle component logistic regression model.

The questionnaire has five main sections that are classified to; personal information, information about residency in Norway, political participation, the surrounded environment and your political view and the last section is the political view. These sections are named as questionnaire class data sets in this thesis.

Fitting the models will be applicable for each data set a side firstly. In section 1, fitting a binomial multiple logistic regression model will be applied for the explanatory variables at each data set aside. The Null hypothesis here is that the explanatory variables has no significant relation with the dependent variable. The alternative hypothesis is; there is a significant relation between the explanatory variables and the dependent variable.

In section 2, principle components analysis will be implemented to each data set aside. Selecting from the principle components that are resulted from the principle components analysis will be based on the eigenvalues. The components that have eigenvalue greater than one will be selected. Furthermore the components that are selected from the principle component analysis in this section will be used to fit a principle component logistic regression model in section 3. The Null hypothesis here is that the principle components has no significant relation with the dependent variable. The alternative hypothesis is; there is a significant relation between the principle components and the dependent variable.

Finding the correlation relation between the significant explanatory variables resulted from each data set in section 1 will be in implemented in section 4. Plus to, gathering all of these

significant explanatory variables to fit a one binomial logistic regression model will be implemented at the same section also. The same hypothesis that are used in section 1 will be followed here as well.

In addition, The correlation relation between the significant principle components resulted in section 7.3 will be defined in section 5. Putting all the significant principle components that are resulted from the principle component logistic regression model fitting in section 3, as an input to fit a new principle component logistic regression one model, will be implemented as well in section 5. The same hypothesis that are used in section 3 will be followed as well in this section.

## **7.1 Results of fitting Logistic regression model**

In this section a binomial logistic regression analysis is applied to each data set of the five data sets, firstly by using voting 2017 national elections as dependent variable. Then repeating again the step of fitting logistic regression model but this time by using voting 2015 local elections variable as the dependent variable.

The type of logistic regression here is binomial multiple logistic regression. The response type is binary. The used model selection method is the stepwise, “forward-backward”, selection technique, that is presented in chapter 6.

### **7.1.1 With the dependent variable: voting 2017 national elections**

- **Personal information**

The explanatory variables are numeric and treated as covariates. No missing values at this data set. After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model has only education status as explanatory variable;

Formula = voting 2017 national elections ~ education status

The significance : No significance

Coefficients table:

	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt; z )</b>
<i>(Intercept)</i>	-1.213	0.913	-1.328	0.1841
<i>Education status</i>	0.375	0.209	1.791	0.073 .

Table (7.1): The logistic regression between voting 2017 and x1 data set variables

The results of fitting the binomial logistic regression are given in table (7.1). The education level is indexed from lower level to higher level. The positive coefficient indicates that the probability of voting in 2017 national elections increases with higher educations.

The predicted probabilities for voting 2017 at each education level are given in Table(7.2)

<i>Education status levels</i>	<b>Level 1</b>	<b>Level 2</b>	<b>Level 3</b>	<b>Level 4</b>	<b>Level 5</b>	<b>Level 6</b>	<b>Level 7</b>
<i>Predicted probability</i>	0.302	0.386	0.478	0.571	0.660	0.739	0.804

Table (0.2):The predicted probabilities for voting 2017 at each education level

- **Information about the residency in Norway (x2)**

The explanatory variables are numeric and treated as covariates. No missing values at this data set. After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model has two explanatory variables;

Formula = voting 2017 national elections ~ residency period +type of residency permit

The significance: with both variables; residency period and type of residency permit.

Coefficients table:

	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt; z )</b>
<i>(Intercept)</i>	-0.7301	1.0674	-0.684	0.4940
<i>Residency period</i>	0.6447	0.2695	2.392	0.0168 *
<i>Type of residency permit</i>	-0.4300	0.2003	-2.147	0.0318 *

Table (7.3):The logistic regression between voting 2017 and x2 data set variables

The results of fitting the binomial logistic regression are given in table (7.3). “Residency period” explanatory variable is indexed from shorter to longer. While “type of residency permit” indexed from stronger to weaker. The positive coefficient indicates that voting 2017 variable increases with the increasing of the “residency period” variable. While the negative coefficient indicates that voting 2017 variable increases with the decreasing of the “residency permit type” variable.

In other words the probability of voting in 2017 elections increases with longer residency period and stronger residency permit.

- **The Political Participation (x3)**

Not all the explanatory variables here are numeric. Six of them are numeric and treated as covariates. While three of them are factors; voting at the original country, why did vote, why didn’t vote. It is converted to numeric to be able of finding correlation and other statistical relations with other explanatory variables. There are missing values at this data set for the factor variables and these missing values are omitted here. The missing values are resulted from the questionnaire structure. After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model has four explanatory variables;

Formula = voting 2017 national elections ~ voting at the original country + signing a petition + why did vote + why didn’t vote

The significance: significance with the variables; “why did vote” and “why didn’t vote”

Coefficients table:

	Estimate	Std. Error	z value	Pr(> z )
<i>(Intercept)</i>	-0.3714	2.1388	-0.174	0.862148
<i>voting at the original country</i>	0.4266	0.2617	1.630	0.103125
<i>Signing a petition</i>	-1.9116	1.1168	-1.712	0.086945
<i>Why did vote</i>	1.4678	0.4199	3.495	0.000473 ***
<i>Why didn't vote</i>	-0.2912	0.1380	-2.110	0.034825 *

Table 0.4) : The logistic regression between voting 2017 and x3 data set variables

The results of fitting the binomial logistic regression are given in Table (7.4). The explanatory variables; "voting at the original country" and "signing a petition" are indexed from higher participation to lower participation. The explanatory variable "why did vote" is indexed from weaker reasons to stronger reasons. Moreover the explanatory variable "why didn't vote" is indexed from weaker reasons to stronger reasons.

The positive coefficients of " voting at the original country " and " why did vote" explanatory variables indicate that the "voting 2017" dependent variable increases with the increasing of them. Which means increasing in the probability of voting in 2017 national election with stronger reason to vote and lower ability of voting at the original country.

While the negative coefficients of "signing a petition" and " why didn't vote" explanatory variables indicate that the "voting 2017" dependent variable decreases with the increasing of them. Which means the probability of voting in 2017 increasing, with the weaker reasons to not vote and the higher participation in signing a petition.

Generally the probability of voting in 2017 elections increases when the reason to vote is stronger, the reason to not vote is weaker, the participation in the elections at home country is lower and the participation in signing a petition is higher.

- **The surrounded environment and your political view (x4)**

All the explanatory variables here are numeric. No missing values at this data set. After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model is;

Formula = voting 2017 national elections ~ 1

The significance: no significance .

- **The Political View (x5)**

The explanatory variables here are 22 ; starting with “e\_po\_15” and continuing to “e\_po\_36”. All the explanatory variables here are numeric. No missing values at this data set.

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model is;

Formula = voting 2017 national election ~ e\_po\_15 + e\_po\_16 + e\_po\_18 + e\_po\_20 + e\_po\_21+ e\_po\_25 + e\_po\_28 + e\_po\_32 + e\_po\_35

The significance: significance with e\_po\_15, e\_po\_18 , e\_po\_20 , e\_po\_25 and e\_po\_32

Coefficients table:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-2.9541	3.0156	-0.980	0.32729
e_po_15	1.0457	0.4045	2.585	0.00974 **
e_po_16	-0.6209	0.3636	-1.708	0.08767
e_po_18	-1.8324	0.6138	-2.985	0.00283 **
e_po_20	2.0592	0.6849	3.007	0.00264 **
e_po_21	-1.0340	0.6036	-1.713	0.08669
e_po_25	-1.2120	0.5377	-2.254	0.02419 *
e_po_28	0.6936	0.3659	1.896	0.05801
e_po_32	1.8413	0.6503	2.832	0.00463 **
e_po_35	-0.3911	0.2466	-1.586	0.11269

Table(7.5):The logistic regression between voting 2017 and x5 data set variables

The results are given in table (7.5). All of the explanatory variables are indexed from strongly disagree (1) to strongly agree (7) . The positive coefficients of “e\_po\_15”,

“e\_po\_20”, “e\_po\_28”, and “e\_po\_32” explanatory variables indicate that the probability of “voting 2017” dependent variable increases with the increasing of them.

It is worth mentioning that these variables refers respectively to the following sentences; “the Norwegian government does not care much about what people like you think”, “voting is a way to make my voice heard on issues I care about”, “politician should focus more on Immigrant women’s needs on their agendas” and “are you financially independent”.

This indicates that the probability of voting in 2017 increases for people who stronger agree with that the Norwegian government does not care much about what people like them “immigrant women” think, but voting is a way to make their voices heard on issues they care about, as politician should focus more on Immigrant women’s needs on their agendas. These people are financially independent.

From the other hand the negative coefficients of “e\_po\_16”, “e\_po\_18”, “e\_po\_21”, “e\_po\_25” and “e\_po\_35” explanatory variables indicate that the probability of voting in 2017 decreases with the increasing of them.

It is worth mentioning that these variables refers respectively to the following sentences; “sometimes politics and government seem so complicated that a person like you cannot really understand what is going on”, “all Norwegians have equal opportunity to vote”, “voting gives me a connection to Norway and other Norwegians”, “I feel that anyone in Norway is able to participate in the political system if they wanted to”. “Traditions affect your political participation”.

This indicates that the probability of voting in 2017 decreases for people who stronger agree with that; all Norwegians have equal opportunity to vote, voting gives them a connection to Norway and other Norwegians and they feel that anyone in Norway is able to participate in the political system if they wanted to. From the other hand these people stronger agree with that traditions affect their political participation and sometimes politics and government seem so complicated that a persons like them cannot really understand what is going on.

Furthermore the last two sentences justify why the probability of voting in 2017 decreases for these people despite they more agree with all Norwegians have equal opportunity to vote, voting gives them a connection to Norway and other Norwegians and they feel that anyone in Norway is able to participate in the political system if they wanted to.

Generally the probability of voting in 2017 increases for people who stronger agree with that the Norwegian government does not care much about what people like them ”immigrant women” think, but voting is a way to make their voices heard on issues they care about, as politician should focus more on Immigrant women’s needs on their agendas. These people are financially independent.

From the other side the probability of voting in 2017 decreases for people who stronger agree with that; all Norwegians have equal opportunity to vote, voting gives them a connection to Norway and other Norwegians and they feel that anyone in Norway is able to participate in the political system if they wanted to. These people stronger agree with that traditions affect their political participation and sometimes politics and government seem so complicated that a persons like them cannot really understand what is going on. And this can be their justification to not vote.

### 7.1.2 With the dependent variable: voting 2015 local elections

- **Personal information (x1)**

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model is;

Formula = voting 2015 local elections ~ Age + social status + education status.

The significance: no significance

Coefficients table:

	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt; z )</b>
<i>(Intercept)</i>	-4.424	1.465	-3.019	0.002**
<i>Age</i>	0.806	0.535	1.506	0.131
<i>Social status</i>	0.354	0.221	1.603	0.108
<i>Education status</i>	0.325	0.202	1.611	0.107

Table (7.6): The logistic regression between voting 2015 and x1 data set variables



The results are given in Table (7.6). The explanatory variable "age" is indexed from younger to older. While the explanatory variable "social status" is indexed from single to divorced and the "education level" is indexed from lower to higher

The positive coefficients indicate that the "voting 2015" dependent variable increases with the increasing of age, social status and education status variables. Which indicates that the probability of voting in 2015 local elections increases for people who are older and having higher education. In comparison with the probability of voting in 2017 national elections. This probability increases only when people have higher education.

- **Information about the residency in Norway (x2)**

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model has two explanatory variables;

Formula = voting 2015 local elections ~ residency period + Norwegian language level

The significance: significance with both variables; residency period and Norwegian language level

Coefficients table:

	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt; z )</b>
<i>(Intercept)</i>	-0.7585	0.9890	-0.767	0.44317
<i>Residency period</i>	0.6228	0.2090	2.980	0.00288 **
<i>Norwegian language level</i>	- 0.5157	0.1976	-2.609	0.00907 **

Table (7.7): The logistic regression between voting 2015 and x2 data set variables

The results are given in Table (7.7). As mentioning before the explanatory variable "residency period" is indexed from shorter to longer. And the explanatory variable "Norwegian language level" is indexed from higher level to lower level.

The positive coefficient of "residency period" explanatory variable indicates that the probability of "voting 2015" variable increases with increasing of "residency period" variable. While the negative coefficient indicates that the "voting 2015" dependent variable

increases as” Norwegian language level “ variable decreases. This indicates that the probability of voting in 2015 local elections increases as” Norwegian language level “ is higher.

Generally the probability of voting in 2015 local elections increases with longer residency period and higher Norwegian language level.

While the probability of voting in 2017 national elections increases with longer residency period and stronger residence permit type.

- **The Political Participation (x3)**

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model is;

Formula = voting 2015 local elections ~ political participation. + voting at the original country + volunteering time for not political party + signing a petition + why did vote + why didn’t vote.

The significance: Significance with political participation, volunteering time for not political party, signing a petition, why did vote, and why didn’t vote

Coefficients table:

	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt; z )</b>
<i>(Intercept)</i>	-3.3268	2.8169	-1.181	0.23759
<i>Political participation</i>	-1.0845	0.4872	-2.226	0.02603 *
<i>voting at the original country</i>	0.4312	0.2968	1.453	0.14630
<i>Volunteering time not to political party</i>	6.5484	2.6191	2.500	0.01241 *
<i>Signing a petition</i>	-4.7821	1.9821	-2.413	0.01584 *
<i>Why did vote</i>	1.7396	0.6238	2.789	0.00529 **
<i>Why didn’t vote</i>	-0.8991	0.3069	-2.930	0.00339 **

Table( 7.8): The logistic regression between voting 2015 and x3 data set variables

The results are given in Table (7.8). The explanatory variable ”Political participation” is indexed from more interested to less interested. The explanatory variables; "voting at the

original country”, “volunteering time not to political party” and “signing a petition” are indexed from higher participation to lower. The explanatory variable “why did vote” is indexed from weaker reasons to stronger reasons. The explanatory variable “why didn’t vote” is indexed from weaker reasons to stronger reasons.

The positive coefficients of “ voting at the original country “ ,” volunteering time not to political party: and “ why did vote” explanatory variables indicate that the “voting 2015” dependent variable increases with the increasing of these variables.

Which indicates that the probability of voting in 2015 local elections increases with lower participation in the elections at the home country , less volunteering time to not a political party and stronger reasons to vote.

While the negative coefficients of “political participation”, “ signing a petition“ and “ why didn’t vote” explanatory variables indicate that the “voting 2015” dependent variable decreases with the increasing of these explanatory variables.

Which indicates that the probability of voting in 2015 local elections increases with more interested in political participation, higher participation in signing a petition and weaker reasons to not vote.

Generally the probability of voting in 2015 local elections increases with lower participation at the elections at the home country , less volunteering time to not a political party, stronger reasons to vote, more interested in political participation, higher participation in signing a petition and weaker reasons to not vote

From the other hand the probability of voting in 2017 elections increases with stronger the reason to vote, weaker reason to not vote, lower participation in elections at home country and higher participation in signing a petition.

- **The surrounded environment and your political view (x4)**

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model is;

Formula = voting 2015 local elections ~ family discuss politics at the past + immigrant friends + pressure on your opinion

The significance: with “family discuss politics at the past”

Coefficients table:

	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt; z )</b>
<i>(Intercept)</i>	0.2573	1.3178	0.195	0.8452
<i>family discuss politics at the past</i>	-0.5518	0.2483	-2.222	0.0263 *
<i>immigrant friends</i>	-0.3216	0.1961	-1.639	0.1011
<i>pressure on your opinion</i>	0.5080	0.3118	1.629	0.1033

Table (7.9): The logistic regression between voting 2015 and x4 data set variables

The results are given in Table (7.9). The explanatory variable “family discuss politics at the past ” is indexed from more often to less often. The explanatory variable “immigrant friends” is indexed from less immigrant friends to more immigrant friends. The explanatory variable “pressure on your opinion” is indexed from larger degree to less degree.

The positive coefficient of “ pressure on your opinion “ explanatory variable indicates that the “voting 2015” dependent variable increases with the increasing of this variables. Which means the probability of voting in 2015 elections increases for less degree of pressure on the participant’s opinion.

While the negative coefficients of “family discuss politics at the past” and “ immigrant friends” explanatory variables indicate that the “voting 2015” dependent variable increasing with the decreasing of these explanatory variables.

This means that the probability of voting in 2015 increases with more often that the family used to discuss politics at the past and with having less immigrant friends.

Generally the probability of voting in 2015 increases with less degree of pressure on the participant’s opinion, more often the family used to discuss politics at the past and having less immigrant friends.

The probability of voting in 2017 national elections doesn’t depend on these explanatory variables.

- **The Political View (x5)**

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model is ;

Formula = voting 2015 local elections ~ e\_po\_18 + e\_po\_19 + e\_po\_20 + e\_po\_22 + e\_po\_24 + e\_po\_27 + e\_po\_32

The significance: significance with e\_po\_19, e\_po\_20, e\_po\_22, e\_po\_24, e\_po\_27, e\_po\_32

Coefficients table:

	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt; z )</b>
<i>(Intercept)</i>	-0.1223	1.5205	-0.080	0.93588
<i>e_po_18</i>	-0.3131	0.1658	-1.888	0.05902
<i>e_po_19</i>	-0.8184	0.2642	-3.097	0.00195 **
<i>e_po_20</i>	0.5941	0.2686	2.212	0.02697 *
<i>e_po_22</i>	-0.3873	0.1955	-1.981	0.04763 *
<i>e_po_24</i>	0.4848	0.2409	2.012	0.04418 *
<i>e_po_27</i>	-0.6507	0.2803	-2.321	0.02027 *
<i>e_po_32</i>	0.4965	0.1942	2.556	0.01058 *

Table (7.10) :The logistic regression between voting 2015 and x5 data set variables

The results are given in table (7.10). All of the explanatory variables are indexed from strongly disagree (1) to strongly agree (7). The positive coefficients of “e\_po\_20”, “e\_po\_24” and “e\_po\_32” explanatory variables indicate that the “voting 2015” dependent variable increases with the increasing of them. It is worth mentioning that these variables refers to the following sentences respectively; “voting is a way to make my voice heard on issues I care about”, “Politicians don’t put Immigrant women’s needs at their agendas” and “Are you financially independent” respectively.

Which indicates that the probability of voting in 2015 local elections increases for people who stronger agree with that voting is a way to make the voice heard on issues they care about, politicians don’t put Immigrant women’s needs at their agendas and they are financially independent.

While the negative coefficients of “e\_po\_18”, “e\_po\_19”, “e\_po\_22” and “e\_po\_27” explanatory variables indicate that the “voting 2015” dependent variable decreases with the increasing of them.

It is worth mentioning that these variables refers to the following sentences; “All Norwegians have equal opportunity to vote”, “It does not make much difference if I vote or not”, “I am satisfied with Norwegian's system of government” and “Immigrant women SHOULD have more influence in politics in Norway” respectively.

Which indicates that the probability of voting in 2015 local elections decreases for people who stronger agree that all Norwegians have equal opportunity to vote, they are satisfied with Norwegian's system of government, Immigrant women SHOULD have more influence in politics in Norway, It does not make much difference if they vote or not.

These results can be explained by; the probability of voting in 2015 local elections decreases for people who stronger agree that all Norwegians have equal opportunity to vote and Immigrant women SHOULD have more influence in politics in Norway. But they believe that it does not make much difference if they vote or not. So they are satisfied with Norwegian's system of government.

Generally, the probability of voting in 2015 local elections increases for people who stronger agree with that voting is a way to make the voice heard on issues they care about, politicians don't put Immigrant women's needs at their agendas and they are financially independent. From the other hand; the probability of voting in 2015 local elections decreases for people who stronger agree that all Norwegians have equal opportunity to vote and Immigrant women SHOULD have more influence in politics in Norway. But it does not make much difference if they vote or not. So they are satisfied with Norwegian's system of government.

At the end of this section; It is worth mentioning that; the whole data set contains 52 explanatory variables. In a trial to fit binomial multiple logistic regression to all of the 52 explanatory variables in one model, a numerical problem occurred and this is might be because of the observations number that are 95 observations, that is not so large comparatively with the number of the explanatory variables.

## 7.2 Principle component Analysis (PCA)

The main goal of this section is to implement principle component analysis method (is presented in chapter 6) to each data set of the five data sets. Since the PCA extracting the most important explanatory variables in a form of components, the aim of PCA here is the dimensions reduction since the data here contains 52 explanatory variables.

Before applying the principle component analysis to each data set a side, the data should be standardised (is presented in chapter 6) and the variables should be numeric, as a preliminary step here to PCA . Data imputation should be applied to the data set that has missing value

Selecting the principle components that are resulted from the principle components analysis here basically will be based on the eigenvalues. The components that have eigenvalue greater than one will be selected<sup>87</sup>. The scree plot and the components' importance table ( that clarifies the data variety amount which each component cover ) will be used to confirm the selection.

- **Personal information (x1)**

The results for the principle component analysis of x1 data set are given in the table of importance of the components table(7.11) , scree plot figure(7.1) and x1 data set components' eigenvalues table(7.12).

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
<i>Standard deviation</i>	1.4424	1.1881	1.1486	0.8464	0.73960	0.71739	0.640
<i>Proportion of Variance</i>	0.2972	0.2016	0.1885	0.1023	0.07814	0.07352	0.05862
<i>Cumulative Proportion</i>	0.2972	0.4989	0.6874	0.7897	0.86786	0.94138	1.00000

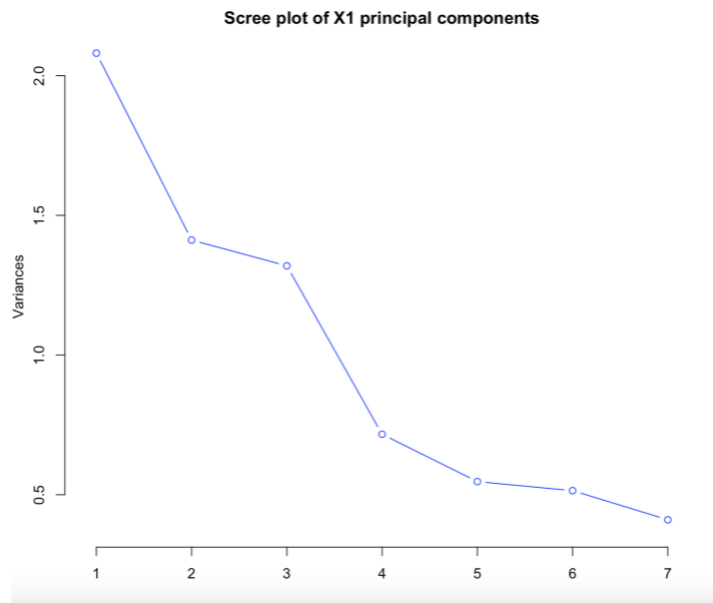
Table (7.11): The principle components characteristics of x1 data set

---

<sup>87</sup> Analytics Vidhya. 2016. "Practical Guide to Principal Component Analysis (PCA) in R & Python."

Table (7.11) shows that the principle components pc1 to pc5 represent around to 86.7% of the data variety. While pc1 to pc4 represent 78.9%. And pc1-pc3 represent 68.7%

The scree plot in general: is a plot which shows the components that represent the maximum variety in the data. The values are represented in a descending order.<sup>88</sup>



Figure( 7.1): Scree plot of x1 data set's principle components

Figure (7.1) shows the scree plot of x1 data set's principle components. The plot shows that pc1,pc2 and pc3 represent the maximum amount of the data variety.

	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>	<b>PC4</b>	<b>PC5</b>	<b>PC6</b>	<b>PC7</b>
<i>Eigenvalues</i>	2.08	1.41	1.31	0.716	0.547	0.514	0.410

Table (7.12): The principle components eigenvalues of x1 data set

Table (7.12) clarifies the x1 components' eigenvalues , it shows that pc1,pc2 and pc3 has eigenvalues greater than 1.

---

<sup>88</sup> Analytics Vidhya. 2016. "Practical Guide to Principal Component Analysis (PCA) in R & Python."



Based on the figure and tables above; pc1,pc2 and pc3 are the best components which represent the maximum data variety, taking into account the best dimensions reduction.

Since the principle components are linear combination of the explanatory variables. Table (7.13) shows the relation between x1 data set explanatory variables and the resulted principle components to this data set. The columns are the principle components and the rows are the explanatory variables. Each value at the table refers to the correlation coefficient between the components and the variables. It shows as well the variety amount which the principle component represents for each explanatory variable in addition

	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>	<b>PC4</b>	<b>PC5</b>	<b>PC6</b>	<b>PC7</b>
<i>Age</i>	0.548	-0.15	0.042	-0.134	0.596	-0.277	-0.472
<i>Social status</i>	0.345	-0.395	-0.435	-0.387	0.084	0.363	0.494
<i>Education status</i>	0.434	0.224	0.422	0.123	-0.136	0.722	-0.162
<i>Employment status</i>	-0.458	-0.215	-0.394	0.200	0.306	0.493	-0.457
<i>Main income</i>	-0.292	0.505	0.089	-0.755	0.248	0.132	-0.045
<i>Family member</i>	0.275	0.313	-0.594	-0.130	-0.526	-0.086	-0.413
<i>Households chores</i>	0.141	0.608	-0.336	0.433	0.430	0.009	0.350

Table (7.13): The principle components' coefficients for x1 data set

For example and based on the table (7.13) above, pc1 has correlation coefficients with “age“= 0.548, “social status“=0.34 and “education status“=0.434 ,

A biplot: is a plot that visualizes how each explanatory variable relates to the principle components and it which direction. <sup>89</sup>

Figure (7.2) shows the biplot that clarifies (x1) explanatory variables positions and directions regarding to pc1 and pc2

---

<sup>89</sup> Analytics Vidhya. 2016. “Practical Guide to Principal Component Analysis (PCA) in R & Python.”

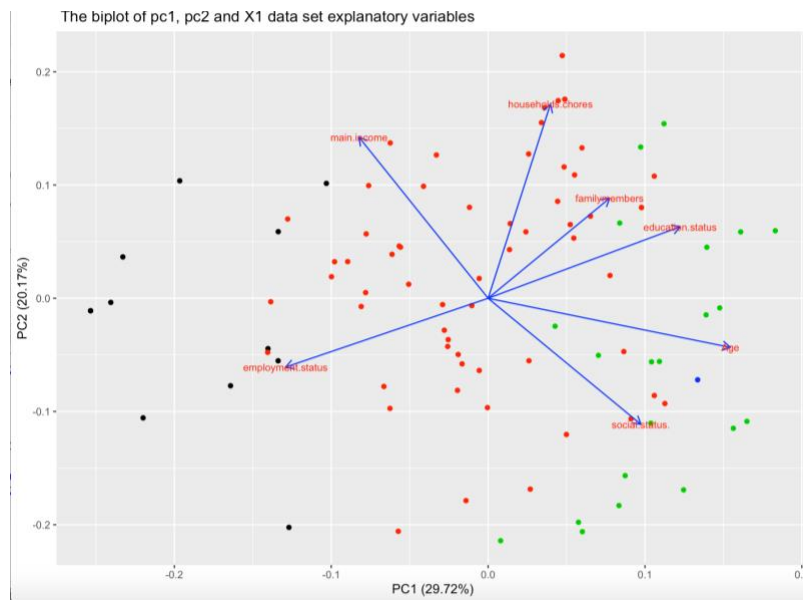


Figure (7.2) : The biplot of pc1, pc2 and the explanatory variables of x1 data set

For example “Age“, “education status“, “social status“ contribute highly to the pc1. High values of these variables move the sample to the right. The colour here based on the age variable distribution in the sample.

- **Information about the residency in Norway (x2)**

The results for the principle component analysis of (x2) data set are given in the components’ importance table(7.14) , scree plot figure(7.3) and the table of x2 data set components’ eigenvalues table(7.15).

	PC1	PC2	PC3	PC4	PC5	PC6
<i>Standard deviation</i>	1.352	1.171	1.060	0.899	0.764	0.529
<i>Proportion of Variance</i>	0.3049	0.228	0.187	0.134	0.097	0.046
<i>Cumulative Proportion</i>	0.3049	0.533	0.721	0.855	0.953	1.00

Table (7.14): The principle components characteristics of x2 data set

Table (7.14) shows that the principle components pc1 to pc5 represent around to 95.3% of the data. While pc1 to pc4 represent 85.5%. And pc1-pc3 represent 72.1%

Figure (7.3) shows the scree plot of x2 data set's principle components. The plot shows that pc1,pc2 and pc3 represent the maximum amount of the data variety.

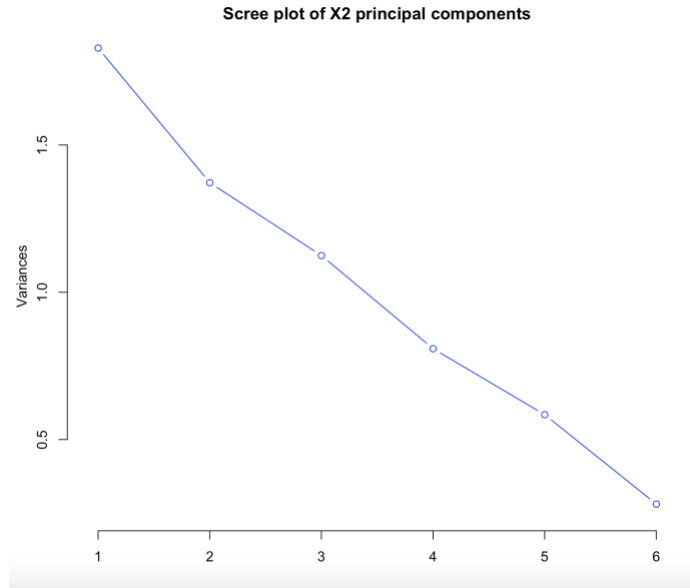


Figure (7.3): Scree plot of x2 data set's principle components

Table (7.15) clarifies the x2 components' eigenvalues , it shows that pc1,pc2 and pc3 has eigenvalues greater than 1.

	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>	<b>PC4</b>	<b>PC5</b>	<b>PC6</b>
<i>Eigenvalues</i>	1.829	1.372	1.124	0.808	0.5844	0.280

Table (7.15): The principle components eigenvalues of x2 data set

Based on the figure and tables above; pc1,pc2 and pc3 are the best components which represent the maximum data variety, taking into account the best dimensions reduction.

Table (7.16) shows the relation between x2 data set explanatory variables and the resulted principle components to this data set.

	PC1	PC2	PC3	PC4	PC5	PC6
<i>Originally from</i>	-0.056	-0.527	-0.551	-0.349	-0.539	-0.003
<i>Status</i>	-0.187	-0.622	0.530	0.092	0.030	-0.535
<i>Residency period</i>	-0.630	-0.249	0.186	-0.026	0.131	0.697
<i>Type of residency permit</i>	0.515	-0.108	0.471	0.176	-0.547	0.411
<i>I live in</i>	0.283	-0.4334	-0.381	0.658	0.356	0.161
<i>Norwegian language level</i>	0.466	-0.268	0.108	-0.635	0.513	0.175

Table (7.16): The principle components' coefficients of the data set x2

Based on the table (7.16) above ,pc1 has correlation with “originally from”= -0.056, “status” = -0.187,” residency period”= - 0.63 , and so on.

Figure (7.4) shows the biplot clarifies the (x2) explanatory variables positions and directions regarding to pc1 and pc2

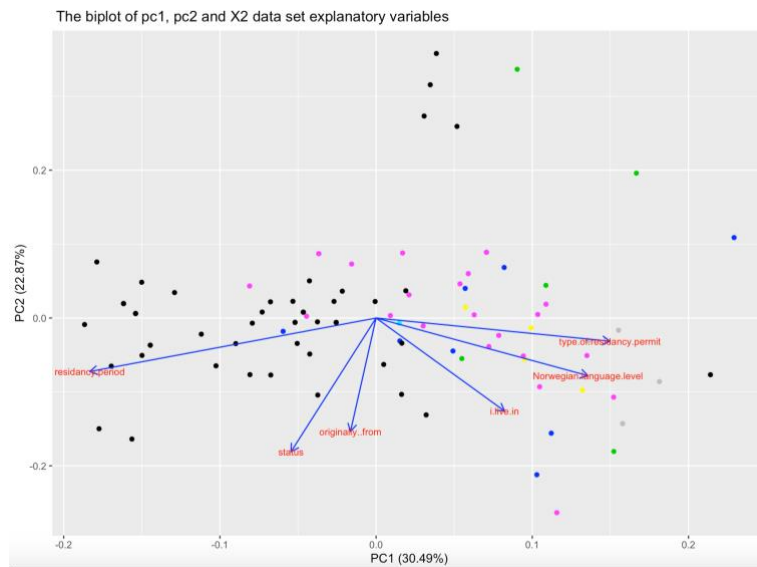


Figure (7.4): The biplot of pc1, pc2 and the explanatory variables of x2 data set

For example “type of residency“ then “Norwegian language level“ contribute highly to the pc1. High values of these variables move the sample to the right. The colour refers to the residency permit type distribution at the sample.

- **The Political Participation (x3)**

The results for the principle component analysis of (x3) data set are given in the table of importance of the components table (7.17) , scree plot figure(7.5) and the table of x3 data set components’ eigenvalues table(7.18).

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
<i>Standard deviation</i>	1.905	1.037	1.006	0.958	0.884	0.767	0.679	0.563	0.461
<i>Proportion of Variance</i>	0.403	0.119	0.112	0.102	0.086	0.065	0.051	0.035	0.023
<i>Cumulative Proportion</i>	0.403	0.522	0.635	0.737	0.824	0.889	0.941	0.976	1.000

Table (7.17): The principle components characteristics of x3 data set.

Table (7.17) shows that the principle components pc1 to pc5 represent around to 82.4% of the data. While pc1 to pc4 represent 73.3 %. And pc1-pc3 represent 63.5%

Figure (7.5) shows the scree plot of x3 data set’s principle components. The plot shows that pc1,pc2 and pc3 represent the maximum amount of the data variety.

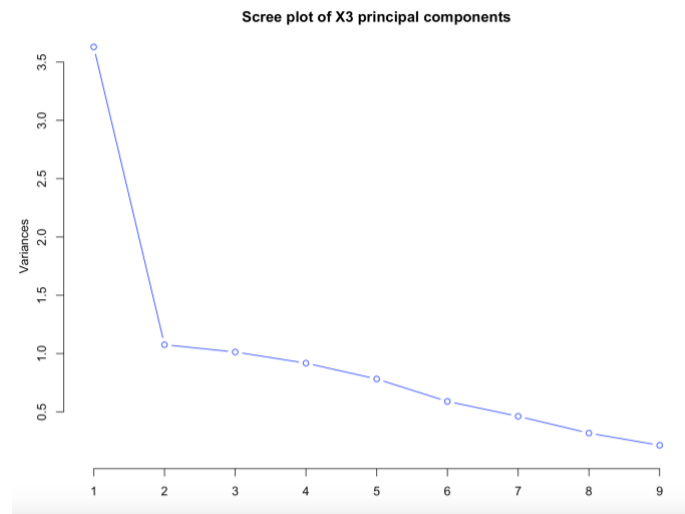


Figure (7.5): Scree plot of x3 data set’s principle components

Table (7.18) clarifies the x3 components' eigenvalues , it shows that pc1,pc2 and pc3 has eigenvalues greater than 1.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
<i>Eigenvalues</i>	3.63	1.075	1.013	0.917	0.781	0.589	0.461	0.317	0.212

Table (7.18): The principle components eigenvalues of x3 data set

Based on the figure and tables above; pc1,pc2 and pc3 are the best components which represent the maximum data variety, taking into account the best dimensions reduction.

Table (7.19) shows the relation between x3 data set explanatory variables and the first six resulted principle components (all are 9 principle components) to this data set.

	PC1	PC2	PC3	PC4	PC5	PC6
<i>Voting at the original country</i>	-0.012	0.810	-0.045	0.525	-0.024	0.239
<i>Member or candidate at political party</i>	0.283	0.001	0.154	-0.069	-0.921	0.088
<i>Member of an interest group</i>	0.406	0.217	-0.183	-0.224	-0.010	-0.296
<i>Volunteering time for non-political party</i>	0.407	-0.213	0.153	0.135	0.070	0.518
<i>Participating in a demonstration</i>	0.444	-0.104	0.151	0.139	0.254	0.258
<i>Signing a petition</i>	0.418	-0.022	-0.266	-0.145	0.204	0.159
<i>Political participation</i>	0.333	0.397	-0.072	-0.418	0.101	-0.264
<i>Why did vote</i>	-0.279	0.086	-0.544	0.437	-0.095	0.598
<i>Why didn't vote</i>	0.160	-0.270	-0.722	0.492	0.137	-0.239

Table (7.19): The principle components' coefficients of data set's x3

Based on the table (7.19) above ,pc1 has correlation with “voting at the original country”= - 0.012, “member or candidate to political party”= 0.283, “member of an interest group”= 0.406, and so on. Figure (7.6) shows the biplot clarifies the (x3) explanatory variables positions and directions regarding to pc1 and pc2

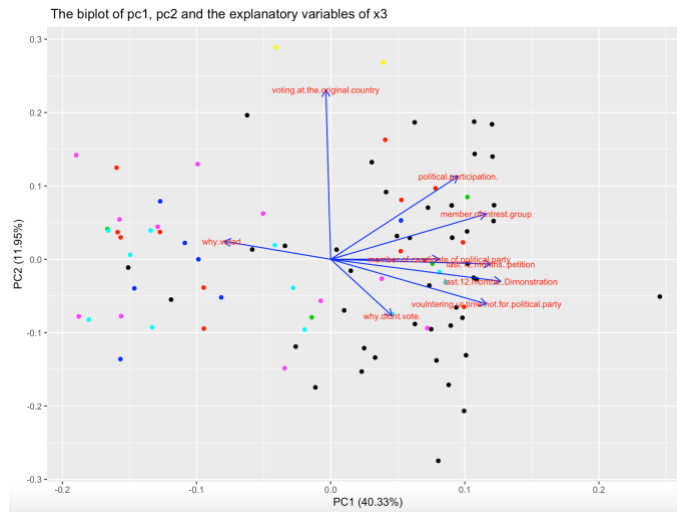


Figure (7.6): The biplot of pc1, pc2 and the explanatory variables of x3 data set

For example “ Participating in a demonstration “ then “ signing a petition“, then ”member of an interest group”, then “volunteering time to not political party”, then” the political participation” contribute highly to the pc1. High values of theses variables move the sample to the right. The colour refers to the “why did vote” variable

- **The surrounded environment and your political view (x4)**

The results for the principle component analysis of (x4) data set are given in the table of importance of the components table(7.20) , scree plot figure(7.7) and the table of x4 data set components’ eigenvalues table(7.21).

	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>	<b>PC4</b>
<i>Standard deviation</i>	1.303	1.040	0.842	0.714
<i>Proportion of Variance</i>	0.424	0.270	0.177	0.127
<i>Cumulative Proportion</i>	0.424	0.695	0.872	1.00

Table (7.20): The principle components characteristics of x4 data set

Table (7.20) shows that the principle components pc1 to pc3 represent around to 87.2% of the data. While pc1to pc2 represent 69.5 %. Figure (7.7) shows the scree plot of x4 data set’s principle components. The plot shows that pc1,pc2 and pc3 represent the maximum amount of the data variety.

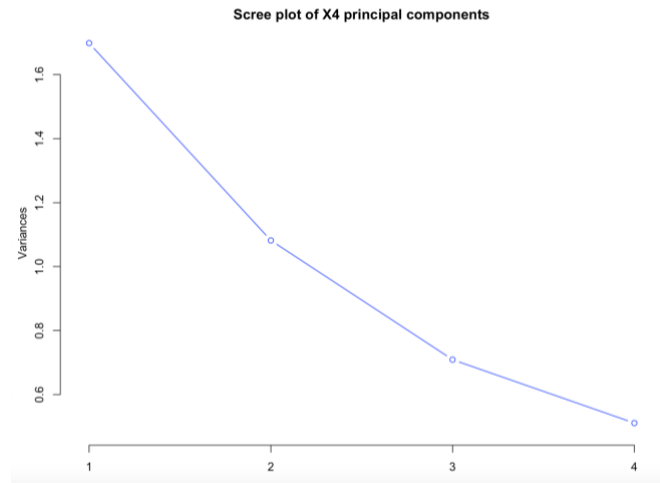


Figure (7.7): Scree plot of x4 data set's principle components

Table (7.21) clarifies the x4 components' eigenvalues, it shows that pc1 and pc2 has eigenvalues greater than 1.

	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>	<b>PC4</b>
<i>Eigenvalues</i>	1.698	1.081	0.708	0.510

Table (7.21): The principle components eigenvalues of x4 data set

Based on the figure and tables above; pc1 and pc2 are the best components which represent the maximum data variety, taking into account the best dimensions reduction.

Table (7.22) shows the relation between x4 data set explanatory variables and the principle components to this data set.

	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>	<b>PC4</b>
<i>Family discussed politics</i>	0.593	-0.271	0.378	-0.656
<i>How often you discuss politics</i>	0.581	0.308	0.406	0.633
<i>Immigrant friends</i>	0.477	0.439	-0.740	-0.177
<i>Pressure on tour opinion</i>	0.284	-0.798	-0.378	0.370

Table (7.22): The principle components' coefficients for the data set x4

Based on the table (7.22) above, pc1 has correlation with "family discussed politics" = 0.593, "how often you discuss politics" = 0.581, "immigrant friends" = 0.477, and so on.



Figure (7.8) shows the biplot clarifies the (x4) explanatory variables positions and directions regarding to pc1 and pc2

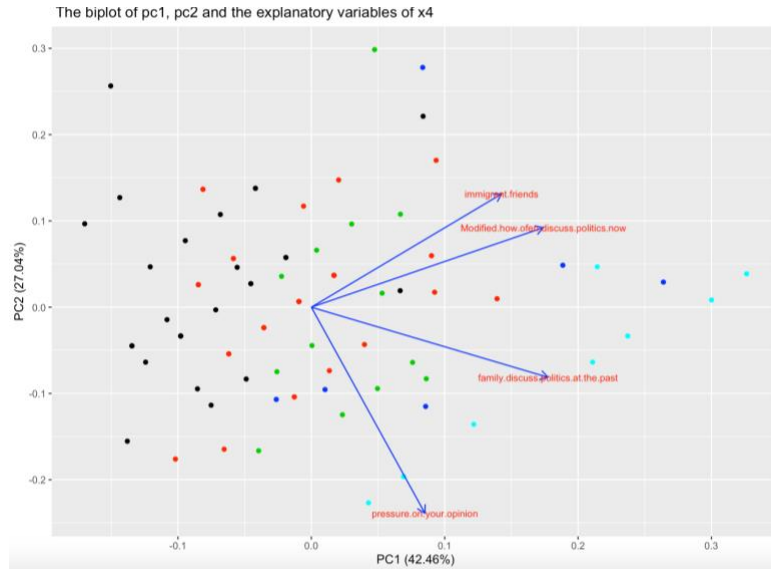


Figure (7.8): The biplot of pc1, pc2 and the explanatory variables of x4 data set

For example “ family discussed politics “ then “ how often you discuss politics“ ,then ”immigrant friends “ variables contribute to highly the pc1. High values of these variables move the sample to the right. The colour is based on the distribution of “family discuss politics at the past”

- **The Political View**

The first eight principle components resulted by the principle component analysis of (x5) data set are given in the table of importance of the components table(7.23) , scree plot figure(7.9) and the table of x5 data set components’ eigenvalues table(7.24).

	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>	<b>PC4</b>	<b>PC5</b>	<b>PC6</b>	<b>PC7</b>	<b>PC8</b>
<i>Standard deviation</i>	1.970	1.775	1.62	1.224	1.153	1.093	1.0418	0.966
<i>Proportion of Variance</i>	0.176	0.143	0.120	0.0681	0.060	0.054	0.049	0.042
<i>Cumulative Proportion</i>	0.176	0.319	0.440	0.5083	0.568	0.623	0.672	0.714

Table (7.23): Some of the principle components characteristics of x5 data set

Table (7.23) shows the first 8 of the principle components (the output principle components are 22). This table shows as well that the principle components pc1 to pc5 represent around to 56.8% of the data. While pc1 to pc6 represent 62.3 %. And pc1-pc7 represent 67.2%.

Figure (7.9) shows the scree plot of x5 data set's principle components. The plot shows that pc1, pc2, pc3, pc4, pc5, pc6 and pc7 represent the maximum amount of the data variety.

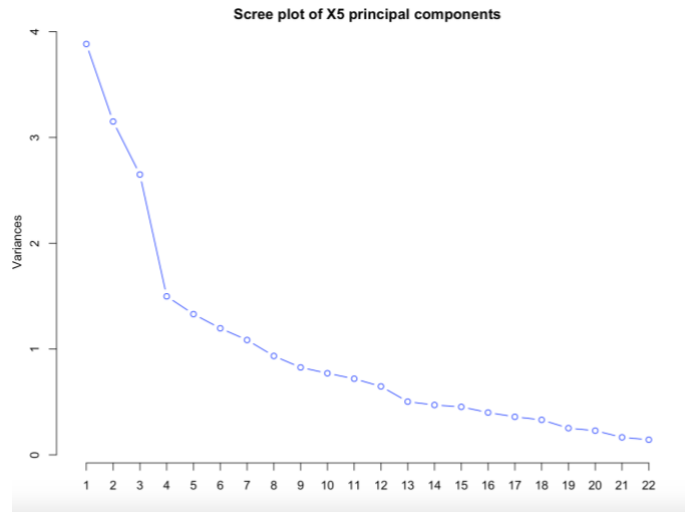


Figure (7.9) : Scree plot of data set's x5 principle components

Table (7.24) clarifies the eigenvalues for the first eight principle components to x5 data set. It shows that pc1 and pc2 has eigenvalues greater than 1.

	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>	<b>PC4</b>	<b>PC5</b>	<b>PC6</b>	<b>PC7</b>	<b>PC8</b>
<i>Eigenvalues</i>	3.88	3.15	2.64	1.49	1.33	1.19	1.08	0.934

Table (7.24): The eigenvalues for some of the principle components for x5 data set

Based on the figure and tables above; pc1,pc2,pc3,pc4,pc5,pc6 and pc7 are the best components which represent the maximum data verity , taking into account the best dimensions reduction. Table (7.25) shows the relation between x5 data set explanatory variables and the first 7 principle components to this data set.

	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>	<b>PC4</b>	<b>PC5</b>	<b>PC6</b>	<b>PC7</b>
<i>e_po_15</i>	0.128	-0.043	0.293	-0.211	0.473	-0.066	0.213
<i>e_po_16</i>	0.294	0.119	0.151	0.098	0.293	-0.116	-0.274
<i>e_po_17</i>	0.194	0.241	-0.140	-0.180	0.184	-0.107	0.075
<i>e_po_18</i>	0.103	0.474	-0.204	-0.566	-0.036	0.004	-0.169
<i>e_po_19</i>	0.150	0.269	0.227	-0.354	0.147	-0.007	-0.054
<i>e_po_20</i>	0.217	-0.262	-0.277	0.177	0.272	-0.053	0.0416
<i>e_po_21</i>	0.213	-0.159	-0.312	0.263	0.238	-0.245	0.171
<i>e_po_22</i>	0.181	0.222	-0.353	-0.068	-0.059	-0.0328	0.243
<i>e_po_23</i>	0.201	-0.202	0.333	-0.002	0.0619	0.0354	0.073
<i>e_po_24</i>	0.295	-0.132	-0.273	-0.17	-0.299	0.366	0.273
<i>e_po_25</i>	-0.436	-0.155	-0.243	0.195	0.295	0.336	0.058
<i>e_po_26</i>	-0.193	0.016	-0.268	0.205	-0.092	-0.219	-0.19
<i>e_po_27</i>	0.0858	0.090	-0.005	0.205	-0.007	-0.170	-0.027
<i>e_po_28</i>	-0.088	0.226	0.047	0.193	-0.028	-0.119	-0.024
<i>e_po_29</i>	0.207	0.364	-0.046	-0.009	0.295	-0.100	-0.126
<i>e_po_30</i>	0.087	-0.317	-0.024	-0.297	0.147	0.099	-0.350
<i>e_po_31</i>	0.070	-0.216	-0.154	0.158	-0.201	0.212	0.036
<i>e_po_32</i>	0.127	-0.181	-0.165	0.265	0.171	-0.084	-0.391
<i>e_po_33</i>	-0.426	0.024	0.510	-0.083	-0.183	-0.129	0.240
<i>e_po_34</i>	0.391	0.386	0.047	0.083	0.102	0.142	-0.052
<i>e_po_35</i>	-0.014	0.121	-0.052	-0.050	-0.295	0.175	0.202
<i>e_po_36</i>	0.184	-0.357	0.549	0.093	-0.125	0.017	-0.273

Table (7.25): The principle components' coefficients for the data set x5

Based on the table (7.25) above ,pc1 has correlation with “*e\_po\_15*” = 0.128, “*e\_po\_16*”= 0.029, “*e\_po\_17*”= 0.194 , and so on. Figure (7.10) shows the biplot clarifies the (x5) explanatory variables positions and directions regarding to pc1 and pc2

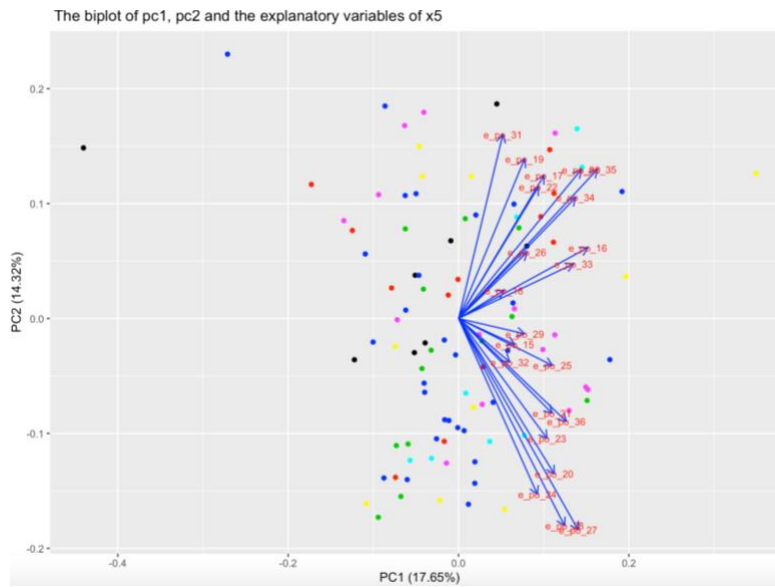


Figure (7.10): The biplot of pc1, pc2 and the explanatory variables of x5 data set

### 7.3 Principle component logistic Regression

The goal of this section is to fit a logistic regression analysis to the selected principle components in the previous section. This analysis called principle component logistic regression (is presented in chapter 6).

The logistic regression is implemented here for each data set aside since the principle components analysis was implemented aside to each data set. This is a complementary step to the previous PCA so it will implemented for each data set twice; once for the dependent variable “voting 2017 ” and the other one is for the variable “voting 2015” as a dependent variables.

#### 7.3.1 With the dependent variable: voting 2017 national elections

- **Personal information**

The selected principle components for x1 data set in section (7.2) are ; pc1,pc2 and pc3.

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model has only pc1;

Formula After using “stepwise” selection: voting 2017 ~ pc1

The significance : No significance

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	0.361	0.268	1.34	0.177
X1pc1	0.334	0.182	1.83	0.0663 .

Table (7.26): The logistic regression between voting 2017 and the x1 data set's selected components

The positive coefficient indicates that the probability of voting in 2017 increases with increasing of pc1. It worth mentioning pc1 here has the positive correlation coefficient with the explanatory variable "age" and "education level". Which means the probability of voting in 2017 increasing for older age and higher education.

- **Information about the residency in Norway**

The selected principle components for x2 data set in section (7.2) are ; pc1, pc2 and pc3.

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model has only pc1;

Formula After using "stepwise" selection: voting 2017 national elections ~ pc1

The significance: with pc1

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	0.231	0.3242	0.713	0.476
X2pc1	-1.125	0.288	-3.900	9.63e-05 ***

Table (7.27): The logistic regression between voting 2017 and the x2 data set's selected components

The negative coefficient indicates that the probability of 2017 voting decreases with increasing of pc1. It worth mentioning that pc1 here has negative correlation coefficient with "residency period" but with negative value=, positive correlation with "residency period permit" and "Norwegian language level".

This indicates the probability of voting in 2017 increases with longer residency period, higher residency permit and higher Norwegian language level.

- **The Political Participation**

The selected principle components for x3 data set in section (7.2) are ; pc1, pc2 and pc3.

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model has pc1 and pc2;

Formula After using “stepwise” selection: voting 2017 national elections ~ pc1+pc2

The significance: with pc1

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	0.802	0.423	1.89	0.058 .
X3pc1	-0.105	0.315	- 3.50	0.0004 ***
X3pc2	0.594	0.328	1.809	0.070 .

Table (7.28): The logistic regression between voting 2017 and the x3 data set’s selected components

The negative coefficient indicates that the probability of 2017 voting variable decreases with increasing of pc1. While the positive coefficient indicates that the probability of 2017 voting variable increases with increasing of pc2

It worth mentioning that pc1 has positive correlation with; participating in a demonstration= 0.444, signing a petition =0.418, volunteering time not for political party= 0.407, member of an interest group=0.406, political participation=0.333, member or a candidate of a political party= 0.28 , why did vote= - 0.279. And the explanatory variables; participating in a demonstration, signing a petition, volunteering time not for political party, member of an interest group, member or a candidate of a political party are indexed from higher participation to lower. Political participation is indexed from more interested to less interested. Why did vote is indexed from weaker to stronger reasons. This means the probability of voting in 2017 increases by higher participation in a demonstration, signing a petition, volunteering time not for political party, member of an interest group, member or a candidate of a political party, more interested in politics and stronger reasons to vote

While pc2 has correlation with; voting at the original country= 0.810 and why didn’t vote = - 0.27. But the explanatory variable voting at the original country is indexed from higher participation to lower . Why didn’t vote is indexed from weaker to stronger. This means that the probability of voting in 2017 increases with lower participation at the elections at the home country and weaker reasons to not vote.

This means that the probability of voting in 2017 increases by higher participation in a demonstration, signing a petition, volunteering time not for political party, member of an interest group, member or a candidate of a political party, more interested in politics stronger reasons to vote, lower participation in the elections at the home country and weaker reasons to not vote.

- **The surrounded environment and your political view**

The selected principle components for x4 data set in section (7.2) are ; pc1 and pc2.

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model has only the intercept, no principle components;

Formula After using “stepwise” selection: voting 2017 national elections ~ 1

The significance: No significance

- **The Political View**

The selected principle components for x5 data set in section (7.2) are ; pc1, pc2, pc3, pc4, pc5, pc6 and pc7. After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model has pc2, pc4, pc5, pc6 and pc7;

Formula After using “stepwise” selection: voting 2017 national elections ~ pc2+ pc4+pc5+ pc6+pc7

The significance: with pc2, pc4, pc5 and pc7

	Estimate	Std. Error	z value	Pr(> z )
<i>(Intercept)</i>	0.481	0.374	1.285	0.198
<i>X5pc2</i>	-0.801	0.235	- 3.40	0.0006 ***
<i>X5pc4</i>	0.733	0.319	2.298	0.0215 *
<i>X5pc5</i>	0.768	0.379	2.024	0.0429 *
<i>X5pc6</i>	0.659	0.3765	1.750	0.0800.
<i>X5pc7</i>	1.105	0.4031	2.743	0.0060 **

Table (7.29): The logistic regression between voting 2017 and x5 data set’s principle components

The negative coefficient indicates that the probability of voting in 2017 decreases with increasing of pc2. While the positive coefficients indicate that the probability of voting in 2017 variable increases with increasing of pc4, pc5, pc6 and pc7. It is worth mentioning that; For pc2 the most important variables are e\_po\_27 and e\_po\_28. The two variables have negative correlation with pc2. These variables refer to the following sentences respectively; “Immigrant women SHOULD have more influence in politics in Norway” and “Politician SHOULD focus more on Immigrant women’s needs on their agendas”. And since pc2 has negative coefficient here, so this indicates that probability of voting in 2017 increases with stronger agree that “Immigrant women SHOULD have more influence in politics in Norway” and “Politician SHOULD focus more on Immigrant women’s needs on their agendas”.

For pc4 the most important variable is e\_po\_19. This variable refers to the sentence “It does not make much difference if I vote or not”. It has negative correlation with pc4, furthermore pc4 has positive coefficient here. Which indicates that the probability of voting in 2017 increases with more disagree with “It does not make much difference if I vote or not”

For pc5 the most important variable is e\_po\_15. This variable refers to the sentence “The Norwegian government does not care much about what people like you think”. It has positive correlation with pc5, furthermore pc5 has positive coefficient here. Which indicates that the probability of voting in 2017 increases with more agree with “The Norwegian government does not care much about what people like you think”

For pc6 the most important variable is e\_po\_32. This variable refers to the sentence “Are you financially independent” It has positive correlation with pc6, furthermore pc6 has positive coefficient here. Which indicates that the probability of voting in 2017 increases with more agree with “Are you financially independent ”

For pc7 the most important variable is e\_po\_26. This variable refers to the sentence “I trust political parties and politicians in Norway”. It has positive correlation with pc7, furthermore pc7 has positive coefficient here. Which indicates that the probability of voting in 2017 increases with more agree with “I trust political parties and politicians in Norway”

Which means, the probability of voting in 2017 increases with more disagree with; “It does not make much difference if I vote or not”, and more agree with “The Norwegian



government does not care much about what people like you think”, “Are you financially independent ”, “I trust political parties and politicians in Norway”, “Immigrant women SHOULD have more influence in politics in Norway” and “Politician SHOULD focus more on Immigrant women’s needs on their agendas”.

### 7.3.2 With the dependent variable: voting 2015 local elections

- **Personal information**

The selected principle components for x1 data set in section (7.2) are ; pc1,pc2 and pc3.

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model has pc1 and pc2;

Formula After using “stepwise” selection: voting 2015 ~ pc1+pc2

The significance: with pc1

	Estimate	Std. Error	z value	Pr(> z )
<i>(Intercept)</i>	-0.271	0.277	-0.976	0.329
<i>X1pc1</i>	0.651	0.230	2.821	0.004 **
<i>X1pc2</i>	-0.309	0.208	- 1.48	0.138

Table (7.30): The logistic regression between voting 2015 and the x1 data set’s selected components

The negative coefficient indicates that the probability of 2015 voting variable decreases with increasing of pc2. While the positive coefficient indicates that the probability of 2015 voting variable increases with increasing of pc1.

Which means the probability of voting in 2015 increasing for older age and higher education. Pc2 has negative correlation with the same variables. And since it has negative coefficient, so this confirms the conclusion at the previous sentence.

- **Information about the residency in Norway**

The selected principle components for x2 data set in section (7.2) are ; pc1,pc2 and pc3.

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model has only pc1;

Formula After using “stepwise” selection: voting 2015 local elections ~ pc1

The significance: with pc1

	Estimate	Std. Error	z value	Pr(> z )
<i>(Intercept)</i>	-0.300	0.28	-1.049	0.29
<i>X2pc1</i>	-0.91	0.23	-3.844	0.0001 ***

Table (7.31): The logistic regression between voting 2015 and the x2 data set’s selected components

The negative coefficient indicates that the probability of voting 2015 increases with decreasing of pc1. This means the probability of voting 2015 increases with increasing of “the residency period variable”, decreasing of “residency permit type” variable, decreasing of “the Norwegian language” variable. So the probability of voting in 2015 increases with longer residency period, stronger residency permit and higher Norwegian language level.

- **The Political Participation**

The selected principle components for x3 data set in section (7.2) are ; pc1,pc2 and pc3.

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model has pc1 and pc3;

Formula After using “stepwise” selection: voting 2015 ~ pc1+pc3

The significance: significance with pc1 and pc3

	Estimate	Std. Error	z value	Pr(> z )
<i>(Intercept)</i>	-0.08	0.308	-0.27	0.78
<i>X3pc1</i>	-0.72	0.179	-4.06	4.9e-05 ***
<i>X3pc3</i>	0.823	0.350	2.34	0.0188 *

Table (7.32): The logistic regression between voting 2015 and the x3 data set’s selected components

The negative coefficient indicates that the probability of 2015 voting variable decreases with increasing of pc1. While the positive coefficient indicates that the probability of 2015 voting variable increases with increasing of pc3.

This means the probability of voting in 2015 increases by higher participation in a demonstration, signing a petition, volunteering time not for political party, member of an interest group, member or a candidate of a political party, more interested in politics, stronger reasons to vote and weaker reasons to not vote.

- **The surrounded environment and your political view**

The selected principle components for x4 data set in section (7.2) are ; pc1 and pc2.

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model has pc1 and pc2;

Formula After using “stepwise” selection:  $\text{voting 2015} \sim \text{pc1} + \text{pc2}$

The significance: significance with pc1

	Estimate	Std. Error	z value	Pr(> z )
<i>(Intercept)</i>	-0.155	0.263	-0.589	0.555
<i>X4pc1</i>	-0.560	0.236	-2.374	0.0176 *
<i>X4pc2</i>	-0.448	0.264	-1.69	0.0894 .

Table (7.33): The logistic regression between voting 2015 and the x4 data set’s selected components

The negative coefficient indicates that the probability of 2015 voting variable decreases with increasing of pc1 and pc2. This indicates that the probability of voting in 2015 increases with more often discussion politics at the past and now, plus to less immigrants friends. And less pressure on the opinion

- **The Political View**

The selected principle components for x5 data set in section (7.2) are ; pc1, pc2, pc3, pc4, pc5, pc6 and pc7. After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model has pc2, pc4, pc5 and pc7;

Formula After using “stepwise” selection:  $\text{voting 2015} \sim \text{pc2} + \text{pc4} + \text{pc5} + \text{pc7}$

The significance: with pc2 and pc7.

	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt; z )</b>
<i>(Intercept)</i>	-0.163	0.288	-0.569	0.569
<i>X5pc2</i>	-0.602	0.184	-3.27	0.001 **
<i>X5pc4</i>	0.419	0.250	1.675	0.093 .
<i>X5pc5</i>	0.468	0.287	1.626	0.103
<i>X5pc7</i>	0.822	0.329	2.49	0.012 *

Table (7.34): The logistic regression between voting 2015 and the x5 data set's selected components

The negative coefficient indicates that the probability of voting in 2015 decreases with increasing of pc2. The positive coefficients indicate that the probability of voting in 2015 increases with increasing of pc4, pc5 and pc7. It is worth mentioning that;

For pc2 the most important variables are e\_po\_27 and e\_po\_28. The two variables have negative correlation with pc2. These variables refer to the following sentences respectively; "Immigrant women SHOULD have more influence in politics in Norway" and "Politician SHOULD focus more on Immigrant women's needs on their agendas". And since pc2 has negative coefficient here, so this indicates that probability of voting in 2015 increases with stronger agree that "Immigrant women SHOULD have more influence in politics in Norway" and "Politician SHOULD focus more on Immigrant women's needs on their agendas".

For pc4 the most important variable is e\_po\_19. This variable refers to the sentence "It does not make much difference if I vote or not". It has negative correlation with pc4, furthermore pc4 has positive coefficient here. Which indicates that the probability of voting in 2015 increases with more disagree with "It does not make much difference if I vote or not"

For pc5 the most important variable is e\_po\_15. This variable refers to the sentence. "The Norwegian government does not care much about what people like you think". It has positive correlation with pc5, furthermore pc5 has positive coefficient here. Which indicates that the probability of voting in 2015 increases with more agree with "The Norwegian government does not care much about what people like you think"

For pc7 the most important variable is e\_po\_26. This variable refers to the sentence "I trust political parties and politicians in Norway". It has positive correlation with pc7, furthermore

pc7 has positive coefficient here. Which indicates that the probability of voting in 2015 increases with more agree with “I trust political parties and politicians in Norway”

This confirms that, the probability of voting in 2015 increases with more disagree with; “It does not make much difference if I vote or not”, and more agree with “The Norwegian government does not care much about what people like you think”, “I trust political parties and politicians in Norway”, “Immigrant women SHOULD have more influence in politics in Norway” and “Politician SHOULD focus more on Immigrant women’s needs on their agendas”

## **7.4 Logistic Regression for all significant explanatory variables**

The main goal of this section is to use the output of section (7.1). Firstly for finding the correlation relation between these significant variables. Then using the significant explanatory variables from each data set and fit a new logistic regression with them twice. One time for the dependent variable “voting in 2017” and the other time for the dependent variable “voting in 2015”. The aim of this section is to check the significance of the significant explanatory variable when putting them together to fit one logistic regression model.

### **7.4.1 Finding the correlation between the significant explanatory variables**

The significant explanatory variables in section (7.1) are; residency period, type of residency permit, Norwegian language level, political participation, signing a petition, volunteering time not to political party, why did vote, why didn’t vote, family discussed politics at the past, e\_po\_15, e\_po\_18, e\_po\_19, e\_po\_20, e\_po\_22, e\_po\_24, e\_po\_25 and e\_po\_32. Some points of the correlation relation is summarized as follows:

- “Residency period” explanatory variable has positive correlation with “why did vote”=0.42 and negative correlation with “type of residency permit”= - 0.411 and “Norwegian language level”=-0.37. Which indicates longer residency period

correlates positively with stronger “why did vote” reason, stronger residency permit type and higher Norwegian language level.

- “Why did vote” has positive correlation with “residency period”= 0.42 and “e\_po\_32”= 0.35 and negative correlation with “ type of residency permit” = - 0.303. Stronger why did vote reason increases with longer residency period , stronger residency permit type and stronger agree with “ I’m financially independent”
- “Why didn’t vote” has negative correlation with “ e\_po\_20” ,” residency period”. Has positive correlation with “e\_po\_19”, “volunteering time not to political party” and “ signing a petition”.

Stronger why didn’t vote reason increases with less participation in “volunteering time to not political party”, less participation in signing a petition, more agree with “It does not make much difference if I vote or not”, shorter residency period and stronger disagree “Voting is a way to make my voice heard on issues I care about

- “e\_po\_18” has positive correlation with “e\_po\_25” and “e\_po\_22”. Which indicates stronger agree with “All Norwegians have equal opportunity to vote” increases with stronger agree with” I feel that anyone in Norway is able to participate in the political system if they wanted to” and “I am satisfied with Norwegian's system of government” .
- “e\_po\_20” has positive correlation with “e\_po\_25”,”e\_po\_32” and negative with “e\_po\_19”. Which indicates stronger agree with “Voting is a way to make my voice heard on issues I care about” increases with stronger agree with” I feel that anyone in Norway is able to participate in the political system if they wanted to” and “I’m financially independent”. It increases with stronger disagree with” It does not make much difference if I vote or not”.

#### **7.4.2 Binomial logistic regression with the dependent variable: voting 2017 national elections**

After trying to fit the binomial multiple logistic regression one model, numerical problems occurred and all  $\Pr(>|z|) = 1$  or .0998. After building the model by adding the variables one by one. The model works fine without one variable, that it is responsible about the numerical

problem, it is the variable “e\_po\_25”. So it is excluded it from the model. So after using the step wise method , the final model is;

Formula = voting 2017 national elections ~ x2\$type of residency permit + x3\$why did vote+ x3\$why didn't vote + x5\$e\_po\_18 + x5\$e\_po\_32

Significance: with type of residency permit, why did vote, why didn't vote and e\_po\_18

	Estimate	Std. Error	z value	Pr(> z )
<i>(Intercept)</i>	3.623	2.816	1.287	0.198
<i>Type of residency permit</i>	-2.012	0.960	-2.09	0.036 *
<i>Why did vote</i>	0.964	0.391	2.46	0.013 *
<i>Why didn't vote</i>	-0.346	0.168	-2.05	0.039*
<i>e_po_18</i>	-1.176	0.593	-1.98	0.047*
<i>e_po_32</i>	1.010	0.576	1.75	0.079 .

Table(7.35): Fitting one logistic regression to all significant variables with voting 2017

The positive coefficients and the negative coefficients indicate that; the probability of voting in 2017 increases with stronger residency permit type, stronger reason to why did vote , weaker reason to why not to vote and stronger agree with “ I'm financially independent”. It increases as well with stronger disagree with “All Norwegians have equal opportunity to vote “, the explanation here depends on the participant's definition to who are these Norwegians. It is interesting that this variable has always negative coefficient, this indicates that participants don't define themselves among those all Norwegian who have equal opportunity to vote

### 7.4.3 Binomial logistic regression with the dependent variable: voting 2015 local elections

After trying to fit the binomial multiple logistic regression one model, numerical problems occurred and all  $Pr(>|z|) = 1$  or .0998. After building the model by adding the variables one by one. The model works fine without one variable, that it is responsible about the numerical problem, it is the variable “e\_po\_24”. So it is excluded it from the model. So after using the step wise method , the final model is;

Formula = voting 2015 local elections ~ Norwegian language level + volunteering time not for political party+ signing petition + why did vote+ why didn't vote + family discussed politics at the past +e\_po\_19

Significance: volunteering time not for political party, signing petition, why did vote+ why didn't vote, e\_po\_19

	Estimate	Std. Error	z value	Pr(> z )
<i>(Intercept)</i>	3.83	3.42	1.11	0.26
<i>Norwegian language level</i>	-0.851	0.464	-1.83	0.066
<i>Volunteering time not for political party</i>	8.75	4.00	2.19	0.028 *
<i>Signing a petition</i>	-7.85	3.61	-2.17	0.029 *
<i>Why did vote</i>	2.39	0.97	2.45	0.014 *
<i>Why didn't vote</i>	-1.06	0.42	-2.51	0.011 *
<i>Family discussed politics</i>	-0.93	0.628	-1.49	0.135
<i>e_po_19</i>	-1.67	0.79	-2.11	0.0343 *

Table(7.36): Fitting one logistic regression to all significant variables with voting 2015

The positive and negative coefficients shows that the probability of voting in 2015 increases with higher Norwegian level, less volunteering time not to political party, higher participation in signing a petition, stronger reasons to why did vote, weaker reasons to why didn't vote, more often family discuss politics and stronger disagree with "It does not make much difference if I vote or not"

## 7.5 Principle component logistic Regression all significant principle components

The main goal of this section is to use the output of section (7.3). Using these significant principle components for each data set and fit a new principle component logistic regression one model with it. One time for the dependent variable "voting in 2017" and the other time for the dependent variable "voting in 2015". The aim of this section is to check the



significance of the significant principle components when putting them together to fit one principle component logistic regression model.

Finding the correlation relation between these significant principle components will be implemented at this section as well.

### **7.5.1 Finding the correlation between the significant principle components**

The significant principle component for the data set (x1) is  $x1_{pc1}$ . The significant principle component for the data set (x2) principle component logistic regression is  $x2_{pc1}$ . The significant principle components for the data set (x3) principle component logistic regression are  $x3_{pc1}$  and  $x3_{pc3}$ . The significant principle component for the data set (x4) is  $x4_{pc1}$ . The significant principle components for the data set (x5) principle component logistic regression are;  $x5_{pc2}$ ,  $x5_{pc4}$ ,  $x5_{pc5}$  and  $x5_{pc7}$ . Some important points of the correlation relation can be summarized here:

- $x2_{pc1}$  : has positive correlation with  $x3_{pc1}$ ,  $x5_{pc2}$ ,  $x3_{pc3}$  and  $x4_{pc1}$ . By looking to the linear combination of each principle components. Here, this correlation relation indicates that: longer residency period correlates positively with stronger reasons why did vote , weaker reason to why didn't vote, higher participation in signing a petition, in demonstration, volunteering time to not political party, member or interest group and more interested in politics and stronger agree with "Immigrant women SHOULD have more influence in politics in Norway"
- $x5_{pc2}$  : has positive correlation with  $x2_{pc1}$ ,  $x3_{pc1}$ ,  $x4_{pc1}$  and negative correlation with  $x5_{pc4}$ ,  $x5_{pc7}$  and  $x5_{pc5}$ . By looking to the linear combination of each principle components. Here, this correlation relation indicates that: stronger agree with "Immigrant women SHOULD have more influence in politics in Norway" and "Politician SHOULD focus more on Immigrant women's needs on their agendas" increase with higher Norwegian language level, stronger type of residency permit, longer residency permit, stronger reasons of why did vote, more interested in politics, higher participation as member in an interest group, higher participation in signing a petition and in a demonstration.

### 7.5.2 With the dependent variable: voting 2017 national elections

No significant principle components for the data set (x1) principle component logistic regression. The significant principle component for the data set (x2) principle component logistic regression is  $x2\$pc1$ . The significant principle component for the data set (x3) principle component logistic regression is  $x3\$pc1$ .

No significant principle components for the data set (x4) principle component logistic regression. The significant principle components for the data set (x5) principle component logistic regression are;  $pc2, pc4, pc5$  and  $pc7$ .

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model is;

Formula = voting 2017 national elections  $\sim x2\$PC1 + x3\$PC1 + x5\$PC4 + x5\$PC5 + x5\$PC7$

Significance :  $x2\$PC1$  and  $x3\$PC1$

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	0.855	0.498	1.716	0.086 .
$X2\$pc1$	-1.012	0.394	- 2.566	0.010 *
$X3\$pc1$	-1.008	0.362	-2.781	0.005 **
$X5\$pc4$	0.643	0.378	1.700	0.089 .
$X5\$pc5$	0.599	0.404	1.483	0.138
$X5\$pc7$	0.693	0.464	1.492	.135

Table (7.37): The logistic regression between voting 2017 and all significant principle components

The positive coefficient indicates that the probability of 2017 voting increases with the increasing of  $x5\$pc4, x5\$pc5$  and  $x5\$pc7$ . The negative coefficient indicates that the probability of 2017 voting variable decreases with the increasing of  $x2\$pc1$ , and  $x3\$pc1$ . It worth mentioning that  $x2\$pc1$  has the highest correlation coefficient with “residency period” but with negative value= -0.63, then with type of “residency period permit “= 0.515 but positive, then “ Norwegian language level”= 0.466. So the probability of voting in 2017

increases for higher Norwegian language level, longer residency period and higher residency permit type.

It worth mentioning that  $x_3$  has correlation with; participating in a demonstration= 0.444, signing a petition =0.418, volunteering time not for political party= 0.407, member of an interest group=0.406, political participation=0.333, member or a candidate of a political party= 0.28 , why did vote= - 0.279.

And the explanatory variables; participating in a demonstration, signing a petition, volunteering time not for political party, member of an interest group, member or a candidate of a political party are indexed from higher participation to lower. Political participation is indexed from more interested to less. Why did vote is indexed from weaker to stronger reasons.

This means the probability of voting in 2017 increases by higher participation in a demonstration, signing a petition, volunteering time not for political party, member of an interest group, member or a candidate of a political party, more interested in politics and stronger reasons to vote

For  $x_4$  the most important variable is  $e_{po\_19}$ . This variable refers to the sentence “It does not make much difference if I vote or not”. It has negative correlation with  $x_4$ , furthermore  $x_4$  has positive coefficient here. Which indicates that the probability of voting in 2017 increases with more disagree with “It does not make much difference if I vote or not”

For  $x_5$  the most important variable is  $e_{po\_15}$ . This variable refers to the sentence “The Norwegian government does not care much about what people like you think”. It has positive correlation with  $x_5$ , furthermore  $x_5$  has positive coefficient here. Which indicates that the probability of voting in 2017 increases with more agree with “The Norwegian government does not care much about what people like you think”

For  $x_7$  the most important variable is  $e_{po\_26}$ . This variable refers to the sentence “I trust political parties and politicians in Norway”. It has positive correlation with  $x_7$ , furthermore  $x_7$  has positive coefficient here. Which indicates that the probability of voting in 2017 increases with more agree with “I trust political parties and politicians in Norway”

### 7.5.3 With the dependent variable: voting 2015 local elections

The significant principle component for the data set (x1) principle component logistic regression is  $x1_{pc1}$ . The significant principle component for the data set (x2) principle component logistic regression is  $x2_{pc1}$ . The significant principle components for the data set (x3) principle component logistic regression is  $x3_{pc1}$  and  $x3_{pc3}$ .

The significant principle component for the data set (x4) principle component logistic regression is  $x4_{pc1}$ .

The significant principle components for the data set (x5) principle component logistic regression are;  $pc2$  and  $pc7$ .

After fitting the binomial multiple logistic regression model using the forward-backward model selection technique, the final model is;

Formula = voting 2015 local elections  $\sim x1_{PC1} + x2_{PC1} + x3_{PC3} + x5_{PC2}$

Significance :  $x1_{PC1} + x2_{PC1} + x3_{PC3} + x5_{PC2}$

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-0.025	0.376	-0.068	0.946
$X1_{pc1}$	0.859	0.338	2.539	0.0111*
$X2_{pc1}$	-0.946	0.349	-2.707	0.006 **
$X3_{pc3}$	1.880	0.530	3.547	0.0003 ***
$X5_{pc2}$	-0.553	0.236	-2.336	0.019 *

Table (7.38): The logistic regression between voting 2015 and all significant principle components

The positive coefficient indicates that the probability of 2015 voting variable increases with the increasing of  $x1_{pc1}$  and  $x3_{pc3}$ . The negative coefficient indicates that the probability of 2015 voting variable decreases with the increasing of  $x2_{pc1}$ , and  $x5_{pc2}$ .

It worth mentioning  $x1_{pc1}$  here has the highest correlation coefficient with the explanatory variable “age”= 0.54, then “education level”=0.43, then “social status”= 0.34. Which means the probability of voting in 2015 increasing for older age and higher education  $x2_{pc1}$  here has the highest correlation coefficient with “residency period” but with negative

value = -0.63, then with type of “residency permit type” = 0.515 but positive, then a positive correlation with “Norwegian language level” = 0.466

Since the probability of voting in 2015 increases with the decreasing of  $x_2$ , this means it increases with increasing of the residency period variable level, decreasing of residency permit type variable level, decreasing of the Norwegian language variable.

But the residency permit type is indexed from higher to lower, the residency period from shorter to longer and the Norwegian level from higher to lower. So the probability of voting in 2015 increases with longer residency period, higher residency permit and higher Norwegian language level.

$x_3$  has correlation with; why didn't vote = -0.722. But the explanatory variable why didn't vote is indexed from weaker to stronger. This means that the probability of voting in 2015 increases with weaker reasons to not vote.

For  $x_5$  the most important variables are  $e_{po\_27}$  and  $e_{po\_28}$ . The two variables have negative correlation with  $x_5$ . These variables refer to the following sentences respectively; “Immigrant women SHOULD have more influence in politics in Norway” and “Politician SHOULD focus more on Immigrant women's needs on their agendas”. And since  $x_5$  has negative coefficient here, so this indicates that probability of voting in 2015 increases with stronger agree that “Immigrant women SHOULD have more influence in politics in Norway” and “Politician SHOULD focus more on Immigrant women's needs on their agendas”.

# Chapter 8

## Conclusion and Discussion

In this thesis a statistical quantitative data collection method was built based on previous existence questionnaires and studies at the same field. The research methodology, questionnaire designing and building were previewed then applied during designing and building the questionnaire structure. This questionnaire was released online. 95 women completed it. And the collected data was used at the data analysis step.

The case study of the statistical study here targets the immigrant women in Norway to learn more about their political participation in Norway by using measuring “ voting in 2017 national elections” and “ voting in 2015 local elections”. The aim of this case study is to address the factors that affect their political participation and the obstacles that challenge their political activities. Plus to, finding the relation between these factors

From the statistical side, different statistical methods were used to explore and analyse the collected data. The data analysis phases ( is presented in details in chapter 7) were divided to 5 phases; taking into account that the data contains 52 explanatory variables and 2 dependent variables. And the number of the completed observations is 95. The questionnaire structure contains 5 main sections. Each section is named as a data set.

The first phase achieved by analysing the data explanatory variables using the binomial logistic regression statistical method. The binomial logistic regression is used here since we have 2 binary dependent variables and this is the ideal method could be applied at this case. In a trial to fit one model with all the 52 explanatory variables , for each dependent variables a side, a numerical and statistical problems occurred and this is in order to the small number of observations (95 records) in a comparison to the number of the explanatory variables. So the logistic regression is applied to each data set a side, furthermore the applied model selection technique was “ the backward forward” method.

The second phase implemented a data dimensional reduction technique using principle components analysis. This technique was implement to each data set aside also. Selection of the resulted principle components was based on the eigenvalues. The third phase used the

selected principle components for each data set to fit a principle components binomial logistic regression model using the “backward forward” model selection technique. Fitting this model was applied to each data set aside for both responses, each one aside.

The fourth phase was achieved by using the significant explanatory variables that were resulted from fitting the binomial logistic regression in the first phase( section 7.1) to each data aside. In this phase the significant variables were used to fit a one new binomial logistic regression model. The correlation between them these variables was defined first.

The fifth phase used the significant principle components that were resulted from fitting the principle components binomial logistic regression in the third phase ( section 7.3) to each data set a side, to fit one new principle components binomial logistic regression model. The correlation between theses principle components was defined at this section as well.

The logistic regression results and the principle components logistic regression are very consistent and give the same findings. But principle components logistic regression added a bit more to the results of the logistic regression, before using principle components analysis(PCA), in a way that confirmed these results and increased the interpretability. And this is expected since PCA is a dimensions reduction technique that and increases the results interpretability<sup>90</sup> and contributes in minimizing the multicollinearity

The main results of logistic regression showed that the probability of voting in 2017 elections increases for people with higher education, longer residency in Norway, stronger residency permission type. Furthermore the probability increases for people who have stronger reasons for voting, weaker reasons for not voting, higher participation in signing a petition at the last 12 months and had a lower participation in the elections at the home country. The probability of voting in 2017 increases for people who are stronger agree that the Norwegian government does not care much about what people like them ”immigrant women” think, but voting is a way to make their voices heard on issues they care about, as politician should focus more on Immigrant women’s needs on their agendas. These people are financially independent.

---

<sup>90</sup> Jolliffe, Ian T., and Jorge Cadima. 2016. “Principal Component Analysis: A Review and Recent Developments.” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*.

The principle components logistic regression confirmed all of these findings and added as well that; voting in 2017 increases for older people, people who have higher participation; in a demonstration, in volunteering time not for political party, as a member of an interest group, as a member or a candidate of a political party and more interested in politics. Furthermore it confirmed as well that probability of voting in 2017 increases for people who are more disagree with; “It does not make much difference if I vote or not”, and more agree with “I trust political parties and politicians in Norway”, “Immigrant women SHOULD have more influence in politics in Norway” and “Politician SHOULD focus more on Immigrant women’s needs on their agendas”.

From the other hand, the results of logistic regression confirmed that the probability of voting in 2015 local elections increases with older age, higher education, moreover longer residency and higher Norwegian language level. It increases with lower participation in the elections at the home country, stronger reasons to vote, more interested in political participation, higher participation in signing a petition the last 12 months and weaker reasons to not vote. Furthermore, it increases with less degree of pressure on the participant’s political opinion, more often the family used to discuss politics at the past and having less immigrant friends. In addition, the probability of voting in 2015 local elections increases for people who stronger agree with that voting is a way to make the voice heard on issues they care about, politicians don’t put Immigrant women’s needs at their agendas and they are financially independent.

The principle components analysis has all these results and added to these results that; the probability of voting in 2015 local elections increases for people with higher Norwegian language level, stronger residency permission type, people who have higher participation in; a demonstration, , volunteering lower time not for political party, higher participation as a member of an interest group, higher participation as a member or a candidate of a political party. It confirms as well that probability of voting in 2015 increases for people who are more disagree with; “It does not make much difference if I vote or not”, and more agree with “I trust political parties and politicians in Norway”, “Immigrant women SHOULD have more influence in politics in Norway” and “Politician SHOULD focus more on Immigrant women’s needs on their agendas”.



It worth confirming that the probability of voting decreases for people who more agree with “Sometimes politics and government seem so complicated that a person like you cannot really understand what is going on” and “Traditions affect your political participation”

Finally the results of logistic regression and principle components logistic regression are very consistent. Both have almost the same findings. Furthermore the principle component logistic regression added a bit to the results of logistic regression in a way of confirming it .

# Bibliography

- Adams, Kimberly S. 2018. "THE SIGNIFICANT INFLUENCE OF PROPORTIONAL REPRESENTATION & DEMOCRATIZATION ON THE SHARE OF FEMALE PARLIAMENTARIANS IN MIDDLE EASTERN AND NORTH AFRICAN (MENA) COUNTRIES." *International Journal of Arts & Sciences*; Cumberland. <https://search-proquest-com.pva.uib.no/docview/2168814533/fulltext/5198A4704A9845BDPQ/1?accountid=8579>.
- Aguilera, Ana M., Manuel Escabias, and Mariano J. Valderrama. 2006. "Using Principal Components for Estimating Logistic Regression with High-Dimensional Multicollinear Data." *Computational Statistics and Data Analysis*.
- Akoglu, Haldun. 2018. "User's Guide to Correlation Coefficients." *Turkish Journal of Emergency Medicine*.
- Analytics Vidhya. 2016. "Practical Guide to Principal Component Analysis (PCA) in R & Python."
- Arunajadai, Sriresh G. 2009. "Stepwise Logistic Regression." *Anesthesia and Analgesia*.
- AVENTURE LINGUISTIQUE. "YOUR LANGUAGE LEVEL." <https://www.aventure-linguistique.ch/en/language-level/>.
- Aware Girls- Commonwealth Foundayion. 2014. Attitudes towards Women's Political Participation-Base Line Survey.
- Ball, Helen L. 2019. "Conducting Online Surveys." *Journal of Human Lactation*.
- Bartram, David. 2016. "Political Integration of Immigrants: Insights from Comparing to Stayers, Not Only to Natives." In *Researching Migrants' Well-Being*,.
- Baştanlar, Yalin, and Mustafa Özuysal. 2014. "Introduction to Machine Learning." *Methods in Molecular Biology*.
- Beckett, Claire, Lennart Eriksson, Erik Johansson, and Conny Wikström. 2017. "Multivariate Data Analysis (MVDA)." In *Pharmaceutical Quality by Design: A Practical Approach*,.
- BENS Research Group MARCS Institute. 2018. *Advances in Principal Component Analysis Research and Developmente*. ed. Ganesh R. Naik. Springer Nature Singapore Pte Ltd.
- Berends, Mark. 2006. "Survey Methods in Educational Research." *Handbook of Complementary Methods in Education Research*.

- Bergh, Johannes, and Tor Bjørklund. 2011. "The Revival of Group Voting: Explaining the Voting Preferences of Immigrants in Norway." *Political Studies*.
- Bermudez, Anastasia. 2016. *Diasporas and Citizenship The Gendered Experiences of Colombian Migrants in Europe* Anastasia Bermudez. eds. Series Editors et al. <https://link-springer-com.pva.uib.no/content/pdf/10.1057%2F978-1-137-53197-1.pdf>.
- Bermudez, Anastasia. 2016. *International Migration, Transnational Politics and Conflict. The Gendered Experiences of Colombian Migrants in Europe*.
- Bertucci, Andrea, David W. Johnson, Roger T. Johnson, and Stella Conte. 2016. "Effect of Task and Goal Interdependence on Achievement, Cooperation, and Support among Elementary School Students." *International Journal of Educational Research*.
- Bilodeau, Antoine. 2016. "Migrating Gender Inequalities? Immigrant Women's Participation in Political Survey Research." *International Migration Review*.
- Bolker, Benjamin M. 2019. "Standard Statistics Revisited." In *Ecological Models and Data in R*.
- Bollard, Stanley S., and James Steve Browder. 2017. "Sampling." In *Practical Handbook of Spectroscopy*.
- Boynton, Petra M., and Trisha Greenhalgh. 2004. "Hands-on Guide to Questionnaire Research: Selecting, Designing, and Developing Your Questionnaire." *British Medical Journal*.
- Bursac, Zoran, Clinton Heath Gauss, David Keith Williams, and David W. Hosmer. 2008. "Purposeful Selection of Variables in Logistic Regression." *Source Code for Biology and Medicine*.
- Chaudhary, Ali R. 2016. *Transnational Politics and Immigrant Political Participation in Europe*.
- Claeskens, Gerda, and Maarten Jansen. 2015. "Model Selection and Model Averaging." In *International Encyclopedia of the Social & Behavioral Sciences: Second Edition*.
- Colosi, Laura. 2006. *Cornell Cooperative Extension Designing an Effective Questionnaire*.
- Community relations, and Directorate of Social and Economic Affairs Council of Europe Press. 1995. *IMMIGRANT WOMEN AND INTEGRATION*.
- Cook, D et al. 2001. "Binary Response and Logistic Regression Analysis." In *Beyond Traditional Statistical Methods*.
- Cooper, Charles J. et al. 2006. "Web-Based Data Collection: Detailed Methods of a Questionnaire and Data Gathering Tool." *Epidemiologic Perspectives and Innovations*.
- Dahlerup, Drude. 2013. *Women, Quotas and Politics* Women, Quotas and Politics.

Dalgard, Anne Berit. 2018. Levekår Blant Norskfødte Med Innvandrereforeldre i Norge 2016.

de Sousa Mendes, Glauco Henrique, and Gilberto Miller Devós Ganga. 2013. "Predicting Success in Product Development: The Application of Principal Component Analysis to Categorical Data and Binomial Logistic Regression." *Journal of Technology Management and Innovation*.

de Souza, Juliana Bottoni, Valdério Anselmo Reisen, Jane Méri Santos, and Glauro Conceição Franco. 2014. "Principal Components and Generalized Linear Modeling in the Correlation between Hospital Admissions and Air Pollution." *Revista de Saude Publica*.

Debbie Wigglesworth , Jalna Hanmer , Catherine Euler, Val Balding, Teresa Ortíz Gómez, Candida Martínez López, Margarita M. Birriel Salcedo, Pilar Ballarín Domingo, Hildur Ve, Berit Bareksten, Nicky Le Feuvre, Eeva Raevaara, Susanna Taskinen. "Women in the European Union." 2015. <http://www.helsinki.fi/science/xantippa/wee/wee24.html>.

Department of Health & Human Services. 2008. "Using Graphs and Charts to Illustrate Quantitative Data." *Evaluation Briefs*.

Dickter, David N. 2012. "Basic Statistical Analysis." In *The Psychology Research Handbook: A Guide for Graduate Students and Research Assistants*.

Dormann, Carsten F. et al. 2013. "Collinearity: A Review of Methods to Deal with It and a Simulation Study Evaluating Their Performance." *Ecography*.

Dray, Stéphane, and Julie Josse. 2015. "Principal Component Analysis with Missing Values: A Comparative Survey of Methods." *Plant Ecology*.

Dudovskiy, John. "Exploratory Research." *research-methodology*. <https://research-methodology.net/research-methodology/research-design/exploratory-research/>.

Ekman, Joakim, and Erik Amnå. 2012. "Political Participation and Civic Engagement: Towards a New Typology." *Human Affairs*.

Escabias, M., A. M. Aguilera, and M. J. Valderrama. 2004. "Principal Component Estimation of Functional Logistic Regression: Discussion of Two Different Approaches." In *Journal of Nonparametric Statistics*.

Etikan, Ilker. 2017. "Sampling and Sampling Methods." *Biometrics & Biostatistics International Journal*.

European Social Survey. 2016. ESS8-2016, Ed.1.0 Study Documentation.

European Union. "Residency Formalities by European Union. -Permanent Residence (>5 Years) for EU Nationals." 2018.

- Farrar, Donald E., and Robert R. Glauber. 1967. "Multicollinearity in Regression Analysis: The Problem Revisited." *The Review of Economics and Statistics*.
- Fielding, Nigel G., Raymond M. Lee, and Grant Blank. 2016. *The SAGE Handbook of Online Research Methods* The SAGE Handbook of Online Research Methods.
- Foddy, William. 1993. *Constructing Questions for Interviews and Questionnaires* Constructing Questions for Interviews and Questionnaires.
- Fossey, Ellie, Carol Harvey, Fiona McDermott, and Larry Davidson. 2002. "Understanding and Evaluating Qualitative Research." *Australian and New Zealand Journal of Psychiatry*.
- Foundations, Statistical. 2008. "Chapter 4 Principle Component and Factor Analysis." *Analysis*.
- Fournier, Patrick, Fred Cutler, Stuart Soroka, and Dietlind Stolle. *Canadian Election Study 2015*.
- Fowler, Floyd J. 2014. *Sage Publications, Inc Survey Research Methods (5th Edition)*.
- Gillies, Duncan Fyfe. 2006. "Lecture 15: Principal Component Analysis." *Intelligent Data Analysis*.
- Given, Lisa. 2008. *The SAGE Encyclopedia of Qualitative Research Methods* The SAGE Encyclopedia of Qualitative Research Methods.
- Grégoire, G. 2015. "Logistic Regression." In *EAS Publications Series*,.
- Hagelund, Anniken. 2002. "Problematizing Culture: Discourses on Integration in Norway." *Journal of International Migration and Integration / Revue de l'integration et de la migration internationale*.
- Haije, Erin Gilliam. 2017. "Top 21 Best Online Survey Software and Questionnaire Tools: An Overview." *Mopinion*. <https://mopinion.com/top-21-best-online-survey-software-and-questionnaire-tools-an-overview/>.
- Hair, Joseph, William Black, Barry Babin, and Rolph Anderson. 2010. "Multivariate Data Analysis: A Global Perspective." In *Multivariate Data Analysis: A Global Perspective*,.
- Harrel Jr., Frank E. 2015. "Regression Modeling Strategies - With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis." *R Software*.
- Hosmer, David, Stanley Lemeshow, and Rodney Sturdivant. 2013. *Applied Logistic Regression*. ProQuest Ebook Central.
- Høstmark, Maria-SSB. 2013. *Velgerundersøkelse Blant Personer Med Innvandrerbakgrunn 2013*.

- Hoveid, Kristin Sandnes, Toril. 2015. Innvandrere Og Norskfødte Med Innvandrerforeldre i et Kjønn- Og Likestillingsperspektiv Utdanning, Arbeid Og Samfunnsdeltakelse.
- Huddleston, Thomas, Jasper Dag Tjaden, and Louise Callier. 2012. How Immigrants Experience Integration in 15 European Cities-Immigrant Citizens Survey.
- Institute Canadian Citizenship. 2015. BALLOTS & BELONGING-NEW CITIZENS ON POLITICAL PARTICIPATION.
- Iversen, Torben, and Frances Rosenbluth. 2010. Women, Work, and Politics: The Political Economy Of Gender Inequality Women, Work, and Politics: The Political Economy of Gender Inequality.
- Jolliffe, Ian T., and Jorge Cadima. 2016. "Principal Component Analysis: A Review and Recent Developments." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*.
- Jolliffe, I T. 2002. "Principal Component Analysis, Second Edition." *Encyclopedia of Statistics in Behavioral Science*.
- Karin Hamre (ed.) et al. 2018. Women and Men in Norway. [https://www.ssb.no/en/befolkning/artikler-og-publikasjoner/\\_attachment/347081?\\_ts=1632b8bcba0](https://www.ssb.no/en/befolkning/artikler-og-publikasjoner/_attachment/347081?_ts=1632b8bcba0).
- Kelley, Kate, Belinda Clark, Vivienne Brown, and John Sitzia. 2003. "Good Practice in the Conduct and Reporting of Survey Research." *International Journal for Quality in Health Care*.
- Kemalbay, Gülder, and Özlem Berak Korkmazoğlu. 2014. "Categorical Principal Component Logistic Regression: A Case Study for Housing Loan Approval." *Procedia - Social and Behavioral Sciences*.
- Khasnabis C, Heinicke Motsch K, Achu K. 2010. *Community-Based Rehabilitation: CBR Guidelines*. Geneva: World Health Organization;
- Kinser, Jason M., and Jason M. Kinser. 2018. "Principle Component Analysis." In *Image Operators*,.
- Kinser, Jason M., and Jason M. Kinser. 2018. "Principle Component Analysis." In *Image Operators*,.
- Kurzman, Charles et al. 2019. "Women's Assessments of Gender Equality." *Sociological Research for a Dynamic World*.
- Laaksonen, Seppo. 2018. *Survey Methodology and Missing Data: Tools and Techniques for Practitioners* Survey Methodology and Missing Data: Tools and Techniques for Practitioners.

- Legan, D., and M. H. Vandeven. 2003. "Sampling Techniques." In *Detecting Pathogens in Food*.
- Li, Yuelin, and Jonathan Baron. 2012. *Behavioral Research Data Analysis with R*.
- Lister, Michael, and Emily Pia. 2008. *Citizenship in Contemporary Europe* *Citizenship in Contemporary Europe*.
- Manning, Christopher. 2007. "Logistic Regression ( with R )." *Changes*.
- Marx, Brian, and Eric Smith. 1990. "Principal Component Estimation for Generalized Linear Regression." Oxford University Press on behalf of Biometrika Trust.
- Mehrjoo, Saeed, and Mahdi Bashiri. 2013. "An Application of Principal Component Analysis and Logistic Regression to Facilitate Production Scheduling Decision Support System: An Automotive Industry Case." *Journal of Industrial Engineering International*.
- Merlo, Juan, Philippe Wagner, Nermin Ghith, and George Leckie. 2016. "An Original Stepwise Multilevel Logistic Regression Analysis of Discriminatory Accuracy: The Case of Neighbourhoods and Health." *PLoS ONE*.
- Moutinho, Luiz, and Graeme Hutcheson. 2011. *The SAGE Dictionary of Quantitative Management Research*. SAGE Publications Ltd. <http://sk.sagepub.com.pva.uib.no/books/the-sage-dictionary-of-quantitative-management-research>.
- Mukaka, M. M. 2012. "Statistics Corner: A Guide to Appropriate Use of Correlation Coefficient in Medical Research." *Malawi Medical Journal*.
- Nardi, Peter. 2006. "Doing Survey Research: A Guide to Quantitative Methods." *Real world research*.
- Nardi, Peter M. 2018. *Doing Survey Research* *Doing Survey Research*.
- Nardi, Peter M., and Peter M. Nardi. 2018. "Developing a Questionnaire." In *Doing Survey Research*.
- O'Neill, Brenda, Lisa Young, Elisabeth Gidengil, and Catherine Côté. 2010. *Women's Political Participation Survey -Quebec-*.
- O'Neill, Brenda, Lisa Young, Elisabeth Gidengil, and Catherine Côté. 2013. "The Political Integration of Immigrant and Visible Minority Women." *Canadian Political Science Review*.
- OECD/European Union (2015). *Indicators of Immigrant Integration 2015*. OECD Publishing, Paris. [https://www.oecd-ilibrary.org/social-issues-migration-health/indicators-of-immigrant-integration-2015-settling-in\\_9789264234024-en](https://www.oecd-ilibrary.org/social-issues-migration-health/indicators-of-immigrant-integration-2015-settling-in_9789264234024-en).

- Ornstein, Michael. 2014. "Designing a Questionnaire." In *A Companion to Survey Research*,.
- Osborne, Jason, and Jason E. King. 2011. "Binary Logistic Regression." In *Best Practices in Quantitative Methods*,.
- Østergaard-Nielsen, Eva. 2003. "The Politics of Migrants' Transnational Political Practices." *International Migration Review*.
- Pak, Son Il, and Tae Ho Oh. 2010. "Correlation and Simple Linear Regression." *Journal of Veterinary Clinics*.
- Paludi, Michele A., and Florence L. Denmark. 2018. *Women and Leadership Women and Leadership*.
- Pamela, Paxton, Melanie M. Hughes, and Jennifer L. Green. 2006. "The International Women's Movement and Women's Political Representation, 1893-2003." *American Sociological Review*.
- Paxton, Pamela, Sheri Kunovich, and Melanie M Hughes. 2007. "Gender in Politics." *Annual Review of Sociology*.
- Peng, Chao Ying Joanne, Kuk Lida Lee, and Gary M. Ingersoll. 2002. "An Introduction to Logistic Regression Analysis and Reporting." *Journal of Educational Research*.
- Peters, Michael A. 2016. "Feminizing World Power: A New Constellation of Women in Politics?" *Journal of Research in Gender Studies*.
- Peterson, Robert. 2014. *Constructing Effective Questionnaires Constructing Effective Questionnaires*.
- Ponto, Julie. 2015. "Understanding and Evaluating Survey Research." *Journal of the advanced practitioner in oncology*.
- Qu, Sandy Q., and John Dumay. 2011. "The Qualitative Research Interview." *Qualitative Research in Accounting and Management*.
- Raykov, Tenko. 2012. *An Introduction to Applied Multivariate Analysis An Introduction to Applied Multivariate Analysis*.
- Roles, Gender, and Gender Identity. 1999. "Gender, Gender Roles, and Gender Identity." *Identity*.
- Roopa, S, and MS Rani. 2012. "Questionnaire Designing for a Survey." *The Journal of Indian Orthodontic Society*.
- Samuel D. Gosling, John A. Johnson. 2010. *Choice Reviews Online Advanced Methods for Conducting Online Behavioral Research*. American Psychological Association.



- Sanguansat, Parinya. 2012. "Two-Dimensional Principal Component Analysis and Its Extensions." In *Principal Component Analysis*,.
- Scott, A. J., D. W. Hosmer, and S. Lemeshow. 1991. "Applied Logistic Regression." *Biometrics*.
- Song, Youngshin, Youn Jung Son, and Doonam Oh. 2015. "Methodological Issues in Questionnaire Design." *Journal of Korean Academy of Nursing*.
- Spangenberg, Eric R., Frederick T. L. Leong, and James T. Austin. 1998. "The Psychology Research Handbook: A Guide for Graduate Students and Research Assistants." *Journal of Marketing Research*.
- Spencer, Neil H. 2013. *Essentials of Multivariate Data Analysis* Essentials of Multivariate Data Analysis.
- Srai, Jaspreet. 2012. *Political Participation amongst Migrants The Case of the Nordic Countries*. UNIVERSITY OF TAMPERE SCHOOL OF MANAGEMENT.
- statista. 2019. *Population in Norway from 2009 to 2019, by Gender*.
- Statistics Norway. 2019. *Lower Growth in the Number of Immigrants*.
- Statistics solutions. "What Is Logistic Regression?"
- Stebbins, Robert. 2012. *Exploratory Research in the Social Sciences* Exploratory Research in the Social Sciences.
- Steinkellner, Alice, and Statistics Norway- Statistisk sentralbyrå(SSB). 2019. *The Growth in the Number of Immigrants in Norway Is the Lowest for a Number of Years. The Increase Was 18 400 in 2018, Making This the Smallest Growth since 2005*.
- Sue, Valerie, and Lois Ritter. 2015. "Planning the Online Survey." In *Conducting Online Surveys*,.
- The National Democratic Institute for international affairs. "WHY WOMEN IN POLITICS?"
- Thomas, Sue. 1996. "Why Gender Matters." *Women & Politics*.
- Tripepi, G., K. J. Jager, F. W. Dekker, and C. Zoccali. 2008. "Linear and Logistic Regression Analysis." *Kidney International*.
- UN.org, and United nations press release. 2005. "NO TOOL FOR DEVELOPMENT MORE POWERFUL THAN WOMEN'S EMPOWERMENT, SAYS SECRETARY-GENERAL, AS WOMEN'S COMMISSION OPENS 2005 SESSION."
- United nations women. "Facts and Figures: Leadership and Political Participation."

- Waltermaurer, Eve. 2012. Questionnaire Design In: Encyclopedia of Epidemiology. ed. Sarah Boslaugh.
- Website, Scholastic. "The History of Women's Suffrage."  
<http://teacher.scholastic.com/activities/suffrage/history.htm>.
- Women, UN. 2016. "Gender Equality – Why It Matters." Sustainable Development Goals.
- Wong, Winnie, and Yves Poisson. 2008. FROM IMMIGRATION TO PARTICIPATION: A Report on Promising Practices in Integration.
- Wright, Kevin B. 2005. "Researching Internet-Based Populations: Advantages and Disadvantages of Online Survey Research, Online Questionnaire Authoring Software Packages, and Web Survey Services." Journal of Computer-Mediated Communication.
- Zapata-Barrero, Ricard, Lorenzo Gabrielli, Elena Sánchez-Montijano, and Thibaut Jaulin. 2013. The Political Participation of Immigrants in Host Countries: An Interpretative Framework from the Perspective of Origin Countries and Societies.
- Zhang, Zhongheng. 2016. "Variable Selection with Stepwise and Best Subset Approaches." Annals of Translational Medicine.

# Appendix A

## Political Participation of immigrant women in Norway during the local elections in 2015 and the national elections 2017

This questionnaire is designed to assess the political participation of immigrant women in Norway during the local elections in 2015 and national elections in 2017. It aims also to collect background information about the respondents , as well as to collect their opinions and views about politics in Norway. The questionnaire is a part of a master thesis in the Data Analysis master program in the Department of Mathematics at the University of Bergen- Norway for the year of 2018/2019. This questionnaire is developed based on questionnaires and call interviews provided by similar studies. This questionnaire is available in Norwegian and English. Chose one language, and answer the questionnaire only once. The targeted group are adult immigrant women in Norway. Your participation will be anonym.

According to the Statistics Norway (SSB) election report in 2013 , immigrant women : *are immigrant women who moved to Norway or women who were born in Norway for both parents with immigrant background (first generation).*

Are you an immigrant woman based on the previous definition : (circle one)

- Yes
- No

Are you 18 years old or older : (circle one)

- Yes
- No

If your answered ‘Yes’ to both questions above , please continue answering the questionnaire. This will take approximately 13 to 15 min. Please answer all questions.

If you have answered “No” to any of the questions above, please stop here.

## Personal information

1. Age : (please circle one)
  - 18- 24
  - 25- 44
  - 45-66
  - 67- 79
  - 80 or above.
  
2. What is your Current Social Status: (please circle one)
  - Single
  - In a relationship
  - Married
  - Widow
  - Divorced or Separated
  
3. What is the highest level of education you have completed ? Please choose the equivalent Norwegian level
  - Didn't complete secondary school
  - Completed Secondary school
  - Completed High school
  - Completed College or technical school diploma
  - Completed University - Bachelor's
  - Completed University - Master's
  - Completed University – PhD
  
4. Which of the following BEST describes your employment situation?
  - Employed full time
  - Employed part time
  - Self-employed
  - Searching for work (unemployed)
  - Caring for family members at home (e.g. children, elderly relatives)
  - Retired
  - Student
  - Working and attending school
  - Other

5. Do you have the main income of the family ?
  - Yes
  - No
  
6. How many family members in your household are you responsible for taking care of (e.g. children, elderly relatives,...) ?
  - None ; I'm responsible only for myself
  - One
  - Two
  - 3-5
  - more than five
  
7. If you live with a partner, how do you divide up household chores?
  - I don't live with a partner.
  - I do most of the chores.
  - My spouse/partner do most of the chores.
  - We share the chores equally.

Information about the residency in Norway:

1. I originally come from :
  - Europe
  - Asia
  - Middle East
  - Africa
  - North America
  - Latin America
  - Australia
  
2. Which of the following does describe your status better?: (please circle one)
  - I was born in Norway, both parents with immigrant background.
  - I moved to Norway .

3. If your answer to the previous question was "I moved to Norway"; please describe when you moved to Norway :

- Less than 5 years ago
- 5-10 years ago
- 11-15 years ago
- 16-20 ago
- more than 20 years ago
- Other.....

4. What type of Norwegian residence permit you have now?

- Norwegian citizen
- Work/entrepreneur
- Study
- Family reunification
- Permit for relatives of an EU/EEA national
- Permanent resident
- Asylum-seeker
- Refugee
- Other

5. I live in:

- Østlandet (Oslo, Akershus, Østfold, Vestfold, Hedmark, Oppland, Buskerud, Telemark )
- Vestlandet (Rogaland, Hordaland, Sogn og Fjordane, Møre og Romsdal )
- Sørlandet (Øst-Agder and Vest-Agder)
- Midt-Norge (Trøndelag )
- Nord-Norge (Nordland, Troms, Finmark, Svalbard)

6. Please describe your Norwegian language level:

- Proficient "First language", C2
- Advanced, C1
- Upper-intermediate, B2
- Intermediate, B1
- Pre-intermediate, A2

- Elementary, A1
- Beginner, A0-A1
- None

### Your Political Participation

1. Are you interested in politics ?
  - Very interested
  - Fairly interested
  - Only Vaguely interested
  - Not interested at all
  - Don't know
  
2. Did you vote in the 2017 national elections?
  - Yes
  - No
  - Not eligible (too young/not a Norwegian Citizen)
  - Refuse to answer
  
3. Did you vote in the 2015 local elections?
  - Yes
  - No
  - Not eligible (too young/ have no permanent residency )
  - Refuse to answer
  
4. If your voted in one or both 2015,2017 elections. Why did you vote? (Please choose the single most important reason), if you didn't vote please skip to question 5.
  - I wanted to support a specific party or candidate
  - I wanted to vote against a specific party or candidate
  - I was concerned about a particular issue or issues raised I think voting is important
  - I wanted to have my voice heard
  - I always vote when I am eligible to vote
  - Family or friends persuaded me to vote
  - Other, please specify \_\_\_\_\_

5. If you didn't vote in one or both 2015,2017 elections Why didn't you voted? Please choose the most important reason.
- Not eligible
  - I didn't know where or when to vote.
  - I didn't think I was on the voters' list.
  - I didn't think my vote would matter to the outcome.
  - I didn't know enough about the candidates, the parties, or the issues.
  - I didn't like the candidates or parties.
  - I wasn't concerned about the issues that were raised.
  - Personal circumstances on election day (e.g. illness, family emergency, out of town). I didn't think I was eligible to vote.
  - I didn't have enough time on election day.
  - I just wasn't interested.
  - I was scared to vote.
  - I didn't speak Norwegian well enough.
  - I don't think any politicians can be trusted.
  - Other, please specify: \_\_\_\_\_

6. If you current status is "I have moved to Norway", before you moved to Norway, did you vote in your country at any elections?
- There were elections and I usually voted
  - There were elections and I sometimes voted
  - There were elections, but I didn't vote
  - There were elections, but I was not eligible to vote
  - There were no elections at all in my country of origin
  - There weren't regular elections.

We are also interested to know more about your political participation in general, please choose the best answer to the following questions.



	Yes	No	Refuse to answer
7. Have you ever been a member or a candidate of a political party?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Have you ever been a member of an interest group working for change on a particular social or political issue?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. In the last 12 months, have you volunteered your time for a group or organization other than a political party?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. In the last 12 months, have you taken a part in a demonstration?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. In the last 12 months, have you signed a petition?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

The surrounded environment and your political view :

1. When you grew up, did your family discuss politics ?
  - frequently
  - occasionally
  - seldom
  - never
  - don't know
  
2. How often do you discuss politics with your family, friends or acquaintance? Would you say that it is :
  - Every day
  - A few times a week

- Never
  - Don't know
3. How many of your friends have an immigrant background?
- None, or almost none
  - Some, but much less than half
  - Around half
  - Much more than half
  - All or almost all
  - don't know
  - refuse to answer
4. To what degree did you experience pressure from your immediate family or close relatives in connection with your choice of a political party
- A large degree
  - A fairly large degree
  - A fairly limited degree
  - A very limited degree
  - NO OPINION

Your Political View :

You are almost done! The following questions ask about your opinion on the Norwegian political system, the sides that affect your political participation in Norway and more about your political participation at the Norwegian society. Please tell us to which extent do you agree with the following statements :

	Strongly disagree	Disagree	Somewhat disagree	Neither disagree nor agree	Somewhat agree	Agree	Strongly agree
1. The Norwegian government does not care much about what people like you think.							
2. Sometimes politics and							

<p>government seem so complicated that a person like you cannot really understand what is going on.</p>							
<p>3. Norwegian government is generally representative of Norwegian's diverse population.</p>							
<p>4. All Norwegians have equal opportunity to vote.</p>							
<p>5. It does not make much difference if I vote or not.</p>							
<p>6. Voting is a way to make my voice heard on issues I care about.</p>							
<p>7. Voting gives me a connection to Norway and other Norwegians</p>							
<p>8. I am satisfied with Norwegian's system of government</p>							
<p>9. Immigrant women don't have enough influence in politics in Norway</p>							
<p>10. Politicians don't put Immigrant women's needs at their agendas</p>							
<p>11. I feel that anyone in Norway is able to participate in the political</p>							

system if they wanted to.							
12. I trust political parties and politicians in Norway.							
13. Immigrant women SHOULD have more influence in politics in Norway							
14. Politician SHOULD focus more on Immigrant women's needs on their agendas							
15. Politics in Norway is for Norwegians							
16. The level of my Norwegian restricts me from voting/participate in Politics							
17. I don't think that my political participation /voting is important							
18. Are you financially independent							
19. Your financial status affects your political participation							
20. Tradition "customs, practices or beliefs that have stood the test of time" is very important to you							

21. Traditions affect your political participation							
22. Do you consider yourself to be a strong supporter to women's rights							

# Appendix B

NO.	The variable description	The variable name
1.	The Norwegian government does not care much about what people like you think.	e_po_15
2.	Sometimes politics and government seem so complicated that a person like you cannot really understand what is going on.	e_po_16
3.	Norwegian government is generally representative of Norwegian's diverse population.	e_po_17
4.	All Norwegians have equal opportunity to vote.	e_po_18
5.	It does not make much difference if I vote or not.	e_po_19
6.	Voting is a way to make my voice heard on issues I care about.	e_po_20
7.	Voting gives me a connection to Norway and other Norwegians	e_po_21
8.	I am satisfied with Norwegian's system of government	e_po_22
9.	Immigrant women don't have enough influence in politics in Norway	e_po_23
10.	Politicians don't put Immigrant women's needs at their agendas	e_po_24
11.	I feel that anyone in Norway is able to participate in the political system if they wanted to.	e_po_25
12.	I trust political parties and politicians in Norway.	e_po_26
13.	Immigrant women SHOULD have more influence in politics in Norway	e_po_27
14.	Politician SHOULD focus more on Immigrant women's needs on their agendas	e_po_28
15.	Politics in Norway is for Norwegians	e_po_29
16.	The level of my Norwegian restricts me from voting/participate in Politics	e_po_30

17.	. I don't think that my political participation /voting is important	e_po_31
18.	Are you financially independent	e_po_32
19.	.Your financial status affects your political participation	e_po_33
20.	Tradition “customs, practices or beliefs that have stood the test of time” is very important to you	e_po_34
21.	.Traditions affect your political participation	e_po_35
22.	Do you consider yourself to be a strong supporter to women's rights	e_po_36

Table(B.1): x5 data set's variables