The use of democratic power by the Swiss population

Business Statistics 1 –Final Project

at the

Faculty of Economics and Business

of the

University of Neuchâtel

with

Prof. Dr. Catalin Starica

as part of the

“Bachelor of Science in Economics”

Authors and student numbers:

Mileci Alba / 21-400-676

Petrini Claire / 23-507-205

Ibrahim Laila / 23-504-897

Comanescu Clara / 23-507-403

Zaepfel Lisa / 23-501-828

Sönmez Mélissa / 19-505-569

Table of Content

Executive Summary

As part of the first year of Business and Descriptive Statistics, we are required to carry out a study on voter turnout in Swiss cantons. We focuse on three socio-economic variables: “linguistic diversity”,“possession of seats in parliament” and “wealth (GDP)”. So, we will analyze Switzerland's 26 cantons in order to verify the impact of these three variables on the exercise of the right to vote.

Our three hypotheses are as follows:

- German-speaking cantons have a higher voter turnout than other cantons.

- Cantons with the highest number of seats in parliament have higher voter turnout than other cantons.

- Wealthier cantons (high GDP) have higher voter turnout than other cantons.

In order to test these hypotheses, we studied the voter turnout with each of these variables using R software.

In other words, we took in consideration the dependence between the participation rate and these three variables.
Using various statistical tools (scatterplots, bar graphs, histograms, density, etc.). We obtained the following results:

In the present situation, the level of participation is high in German-speaking cantons with a small surface area, i.e. when a large majority of people in the canton take part in the vote.

Indeed, it seems that the smaller the canton, the more people vote. We therefore believe that voter turnout is closely linked to the size of a canton.

We also note that a canton with a large population has more seats in parliament than cantons with smaller populations.
So, the latter do not have a high participation rate. We assume that the smaller the canton, the fewer seats it has in parliament, and the more it will vote.

For the last hypothesis, we can argue that cantons with less wealth (low GDP) would have higher voter turnout.

There should be no generalization because it is not always the case.

1. Introduction

With the intention of better understanding Switzerland, we will carry out a statistical study on the use of democratic power by the Swiss population.

First of all, in order to better define the research framework, the population chosen is the 26 cantons of Switzerland. Taking the cantons population, as a sample, is interesting because they are independent States. Thus, the cantons are part of the federal system by which Switzerland is governed, including the Confederation and the municipalities.

Secondly, Switzerland is a direct democracy, i.e. the population can participate in all political phases. It is highly relevant to study a country with cultural, geographical and linguistic diversity, as it enables us to relate the linguistic variable and the participation of cantons in democratic power.

Finally, to avoid confusion in the analysis, it is necessary to restrict the study. The research themes have been chosen to cover several socio-economic fields. Indeed, this study of Swiss democratic power will attempt to statistically analyze the following three hypotheses:

♦ German-speaking cantons have a higher voter turnout than other cantons.

♦ Cantons with the highest number of seats in parliament have higher voter turnout than other cantons.

♦ Wealthier cantons (high GDP) have higher voter turnout than other cantons.

2. Data

We decided to focus our research on Switzerland, this includes all 26 cantons. However, our data does not take into account the fact that there are bilingual and quadrilingual cantons, as well as Romansh.

In terms of voter turnout, we will use the rate that ranks cantons by their level of participation from 1 to 100 (1 corresponding to the canton with the lowest participation, and 100 indicating the canton with the highest participation). Each canton's participation rate will be analyzed according to three variables: linguistic diversity, possession of seats in parliament and wealth (GDP).

First of all, we need to define what these variables are:

★ **Linguistic diversity** refers to the fact that Switzerland is a country with four official national languages (German, French, Italian and Romansh). German is the most widely spoken language, followed by French, Italian and Romansh.

★ **The possession of seats** in parliament represents the number of seats a canton has in parliament. In parliament, there are two chambers: the National Council and the Council of States. In the National Council, seats are allocated according to the size of the cantons' populations, so the more inhabitants a canton has, the more seats it has. The most populous canton will have 36 seats, and the least populous only one. In the Council of States, on the other hand, all 20 cantons have two seats, but the half-cantons of Obwalden, Nidwalden, Basel-Stadt, Basel-Landschaft, Appenzell Innerrhoden and Appenzell Ausserrhoden have just one seat.

★ **GDP** measures the wealth generated by the cantons. This indicator provides information on cantonal economic activity in Switzerland. The richest cantons are Zurich, Bern, Vaud and Geneva.

2.1. Data gathering

We have collected all the data thanks to the official website of the Swiss Confederation "admin" (bfs.admin.ch and pxweb.bfs.admin.ch). Afterwards, we chose a specific year (2021) which will be the basis of our research. Moreover the data do not vary greatly and remain consistent with all our informations.

Based on this information, we wanted to find out whether cantonal voting rates could help us understand why the voting rate in Switzerland is higher or lower depending on the language, wealth and number of seats in parliament of our 26 cantons.
So, cantonal voter turnout will be entered as percentages, and for the variables, as already mentioned, we have chosen three that will be fixed.

So we'll start with the voting rates of the cantons that will be our "y", and add some variables, our "x", which will be:

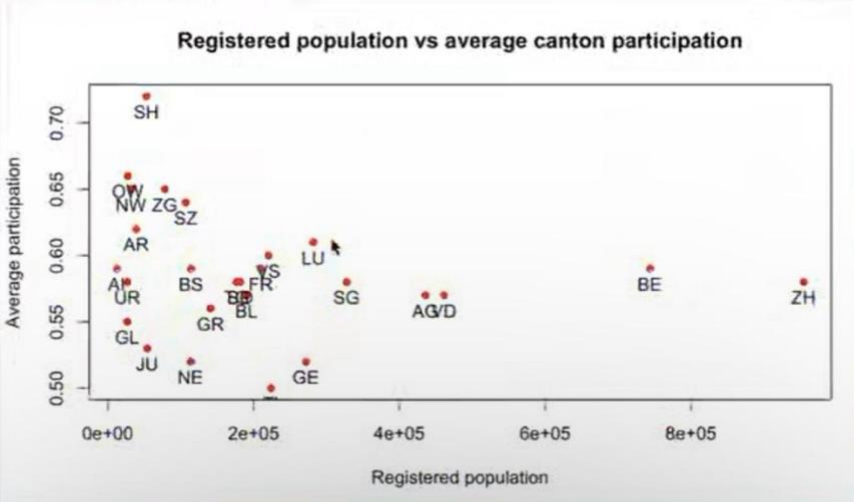
• The different languages of the cantons (German-speaking, French-speaking and Italian-speaking)

• The number of seats that cantons have in Parliament

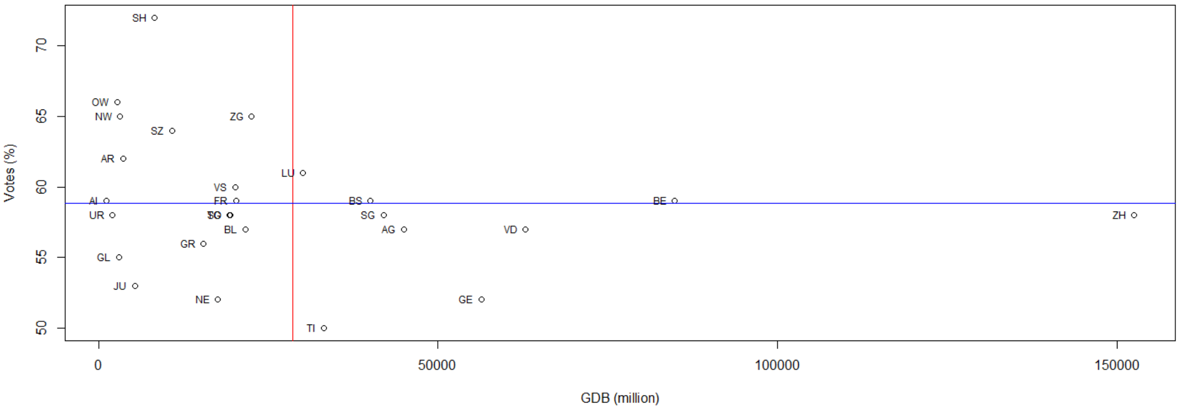
• The wealth of the cantons (GDP)

2.2. Descriptive overview

Here is our scatter plot, which shows the relationship between each canton's participation rate and its registered population.



This second scatterplot allows us to vizualize the relationship and the impact of each canton's GDP on its voting percentage.



His analysis in more details will confirm or not our hypothesis concerning GDP.

This gives us a global picture of this relationship, but also, the participation rate variable which will be the focus of our analysis.

3. Methodology

To do this, we will use a statistical analysis software package, which is called R, to analyze various aspects of our data. The aim is to identify relationships between voter turnout and our variables. We will carry out a statistical and systematic analysis of the three spheres expressed by our variables.

Our study is divided into three parts of analysis, representing the parts of the three variables. In our analysis, we will also introduce the term of density.

Moreover, each variable will be associated with the participation data in order to verify the dependence between the two variables. Scatterplots, bar graphs and histograms will be the main tools used.

Finally, we will seek to answer our assumption through a relevant discussion on our project.

4. Analysis

4.1. Swiss national languages

The language variable represents Switzerland's three main languages: German, French and Italian.
This variable is a quantitative one, indicating the rate of participation a canton can have.
Moreover, this variable is discrete, as it is possible to enumerate its values.

To start with, here's a scatter plot. This type of graph is a good tool for analysis.
analysis tool, placing each canton on the graph according to its participation rate represented on the y-axis (Y) and their population on the x-axis (X). To facilitate analysis, the median of the two variables has been plotted on both axes.

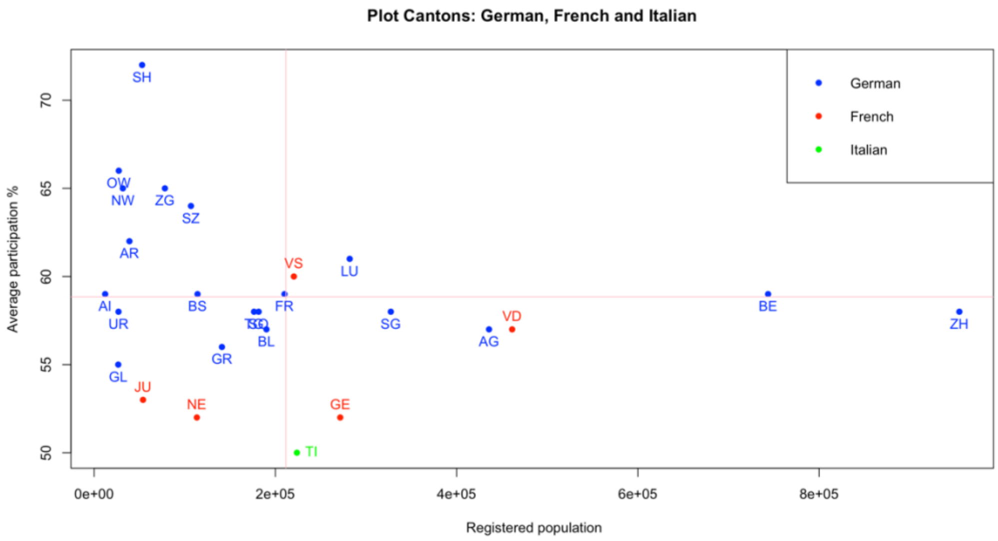


Figure 1 : plot cantons (German, French and Italian) graph.

First, we can see on this graph that the canton with the largest population is Zürich (ZH), its average turnout is below 58%, so it's below the median.
And the canton with the highest average turnout is Schaffhausen (SH).
over 70%. We can say that the German-speaking cantons are at the extremes of each axis
of each axis (highest participation and highest population). As a result, German-speaking cantons with large populations have a low average turnout (ZH, BE, AG, SG), while those with smaller populations have a higher average turnout (SH, OW, NW, ZG, SZ).

Secondly, among the French-speaking cantons, the canton of Vaud (VD) has a larger population and a turnout rate of between 56%-57%. The French-speaking canton with the highest participation rate is Valais (VS), with a rate of 60%. So, we can explain these results by the fact that the French-speaking cantons have smaller populations, but have turnout rates in excess of 50%.

And thirdly, the only Italian-speaking canton (TI) has a low population and an equally low turnout rate of 50%. This is the canton with the lowest turnout.

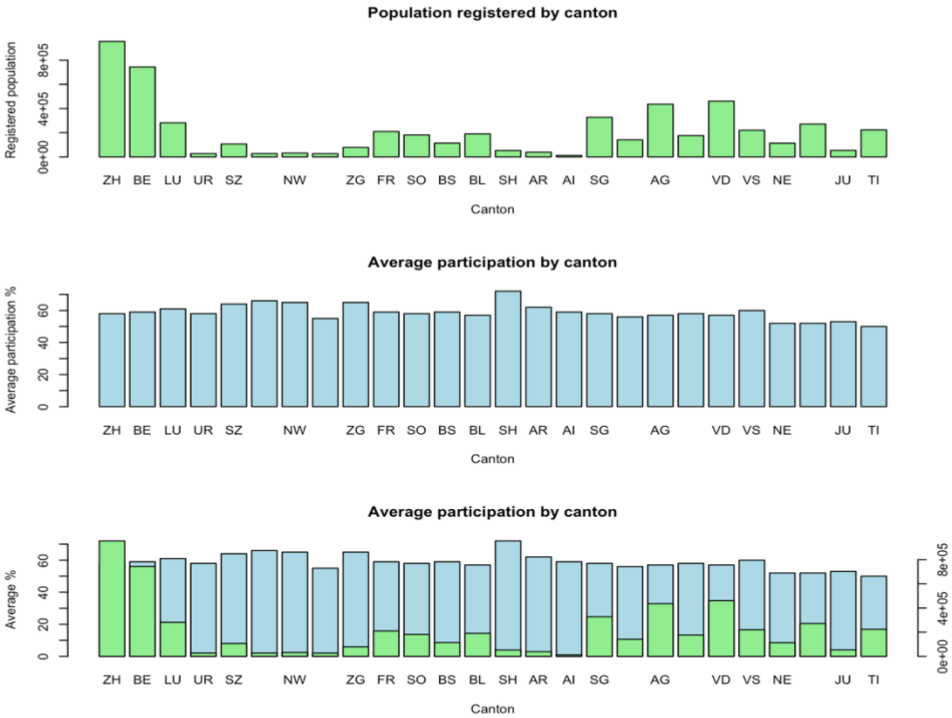


Figure 2 : Average participation by canton graph.

Next, we drew three histograms. This type of graphical representation makes it possible to Put in reéation different variables, such as population per canton and average of the voter turnout per canton. The German-speaking cantons are on the left, followed by the French-speaking cantons and, at the end, the only Italian-speaking canton.

The first graph shows the registered population of each canton in Switzerland. We can see that there are major differences between the two German-speaking cantons of Zürich (ZH) and Berne (BE) and the rest of Switzerland. These two cantons are the most populous in the country.

The second graph shows the average turnout of the population at the polls. We can see that the cantons of Zürich (ZH) and Berne (BE), being the most populous, are not the cantons with the highest turnout rates. The cantons with the highest rates are Schaffhausen (SH) and Obwald (OW), both are among Switzerland's least populous cantons.

And the last graph represents the union of the first two.
Finally, thanks to this, we can see the differences between the registered populations and the average of the voter turnout for each canton. And, in this way, we can see the difference between linguistically different cantons more easily.

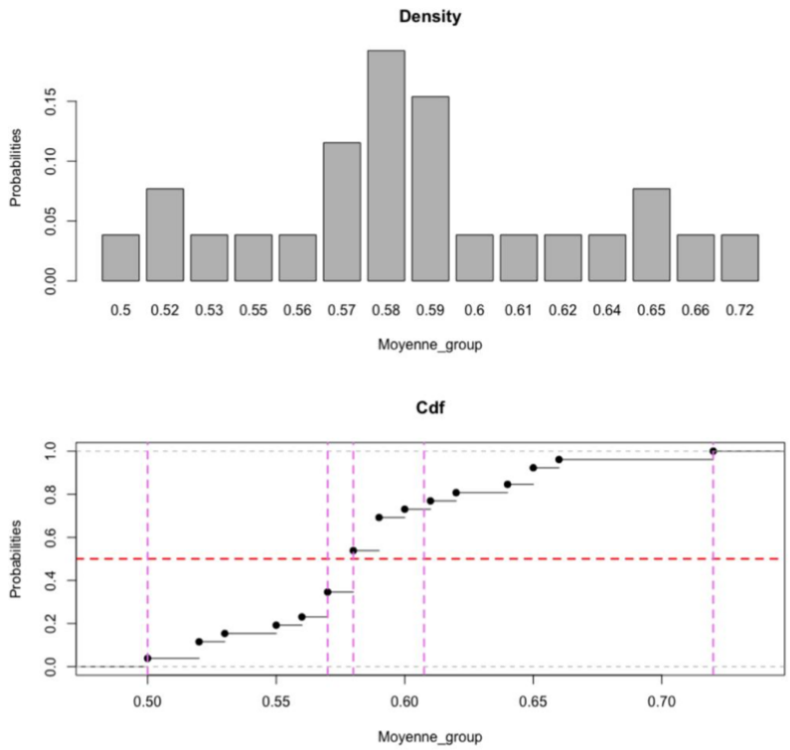


Figure 3 : Density and Cdf graphs

Then,the bar graph then highlights the difference in density of the different participation rates. The CDF shows us more clearly the jumps in average of participation for all cantons. The pink lines represent the quantile we found thanks to the median of 58%, represented by the red line.

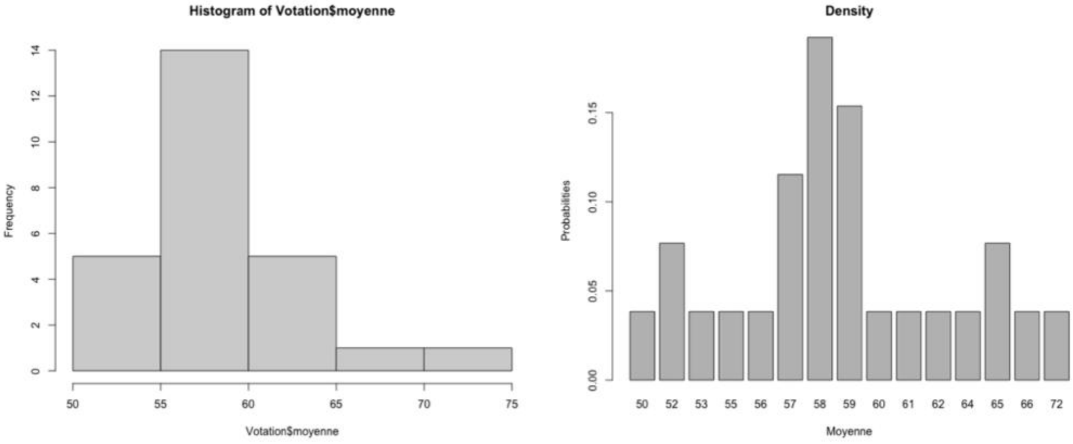


Figure 4 : Histogram and Density

This histogram shows us the frequency of the average of voter turnout of cantons, i.e. the frequency of cantons between 50% and 55%, those between 55% and 60%, and so on.
Thus, we can say that the highest frequency of average of voter turnout is found at a frequency of 14, with cantons having an average voter turnout between 55% and 60%.

The second graph of density shows the population of the various cantons, according to their voter turnout. As we saw earlier, the canton with the largest population is Zürich (ZH) (as its "bar" exceeds 0.15), with a turnout rate of around 58%, followed by Berne (BE).

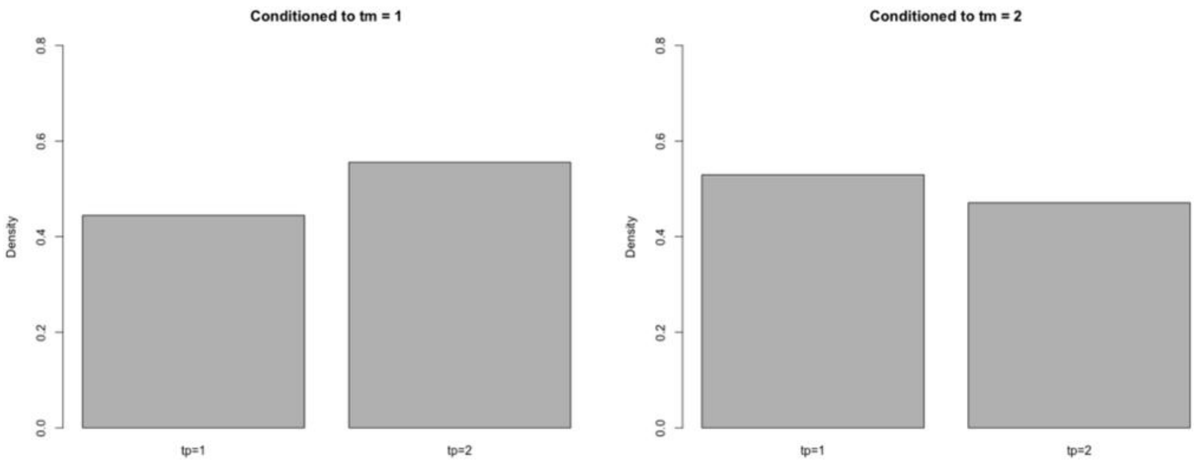


Figure 5: Conditioned 1 and 2.

The graph on the left shows the participation rate, which appears to be low (tm = 1), with "tp = 1" representing cantons with small populations and "tp = 2" representing cantons with large populations. We can conclude that low participation corresponds to cantons with a large registered population.

The graph on the right shows the high participation rate (tm = 2).
We can see that cantons with a smaller registered population have a higher voter turnout.

Conclusion:

Finally, with all the elements at our disposal, the initial hypothesis, according to which German-speaking cantons participate more than other cantons, seems to be verified.

Indeed, it seems that cantons with small populations participate more than cantons with large populations. As a result, German-speaking cantons with small populations would participate more than German-speaking cantons with large populations. It would then be possible to consider that the linguistically mixed cantons that would participate the most are those with the smallest populations.

4.2. Seats Swiss Parlement

As a second variable, we chose the number of seats each canton has in Parliament. This way, we want to analyse the influence of the number of seats each canton has, on the participation of thes cantons in votes, and to see how important is this influence.

This is a quantitative variable, as the different values of the variable are quantities. It is also discrete, since it takes only a finite number of values.

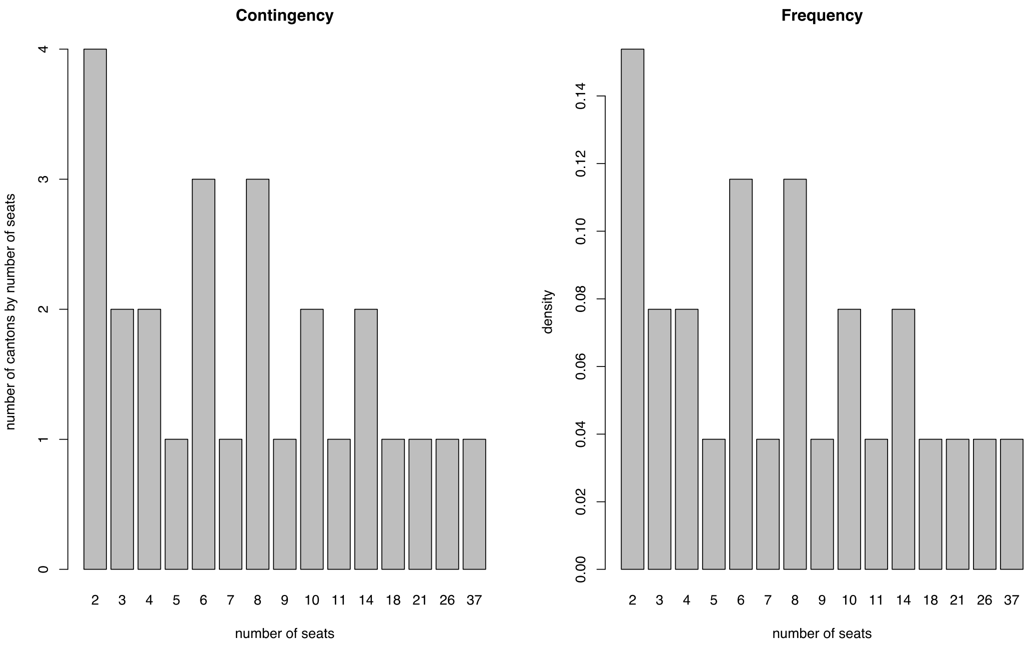


Figure 6 : Contingency and Frequency

Figure 6 shows a contingency table and a frequency table. In the contingency table, the ordinate shows the number of cantons with a certain number of seats, while the abscissa shows this number of seats. In the frequency chart, we can observe the density, i.e. the probability of obtaining a certain number of seats in all our cantons.

We can see that the minimum number of total seats in parliament is 2. This number is the highest (the bar is the highest in the graph) among our cantons, and therefore the probability in our frequency table is the highest.

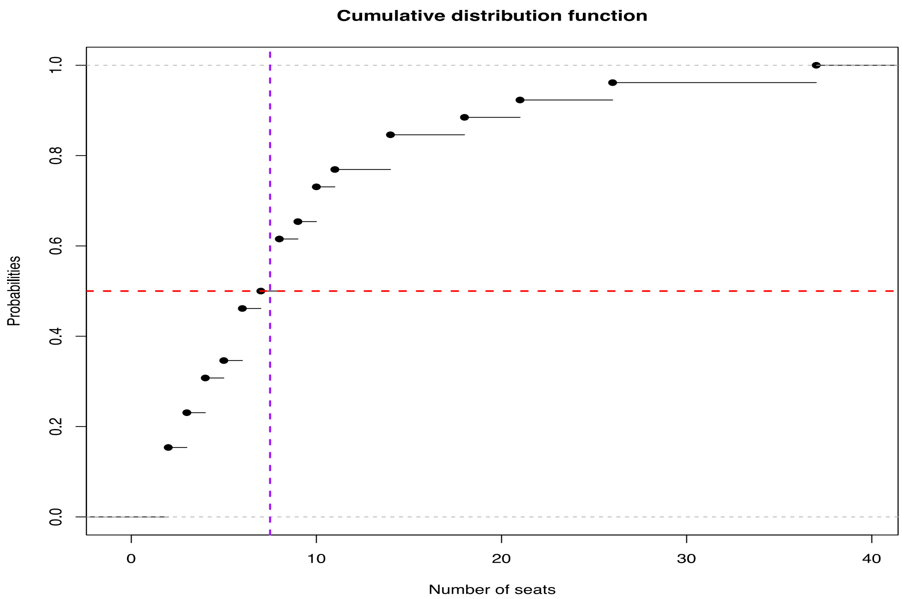


Figure 7 : Cdf (number of seats and probabilities) graph

This graph shows the distribution function of our variable. We can see the possible values of our variable as a function of cumulative frequencies. The possible values of our variable X (Number of seats) are highlighted by the discontinuities, i.e. the different jumps from one bar to the next.

The median, which shows the number of seats for which, below 50% of cantons with the fewest seats are and above 50% of cantons with the most seats are, is 7.5 (found more precisely using the quantile function). Thus, the first 50% of cantons have at least 7.5 seats in parliament.

This graph shows that the majority of cantons have a small number of seats in parliament, generally fewer than 10.

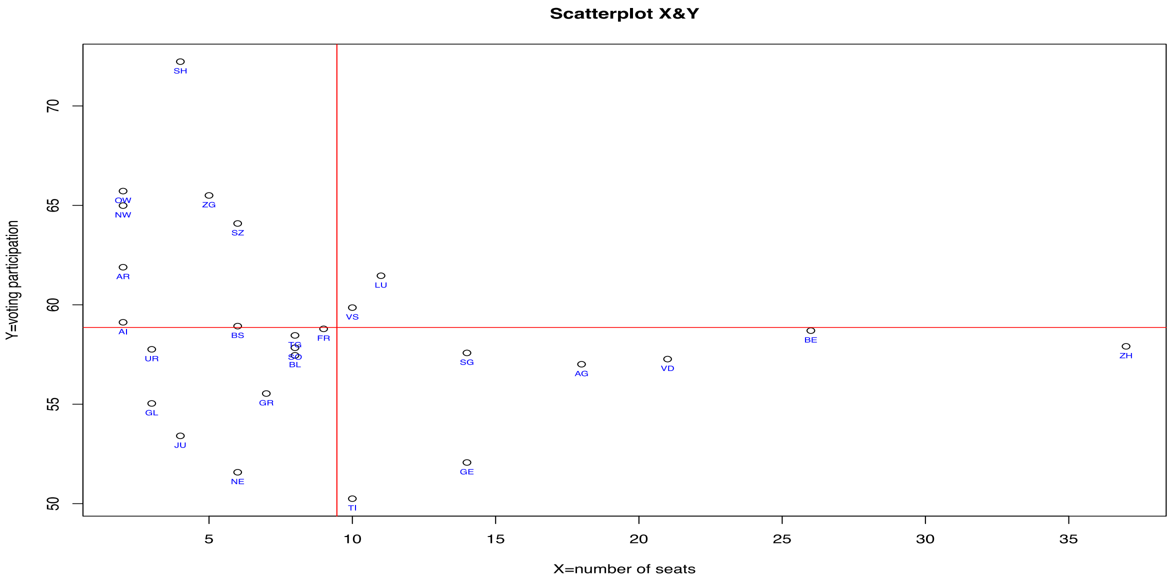


Figure 8: Scatterplot X&Y

This scatterplot shows the relationship between the number of seats in parliament, represented on the x-axis, and voter turnout, represented on the y-axis. It's important to note that our Y variable is evaluated as a percentage. To make the graph easier to understand, we have added the abbreviations of the corresponding cantons for each point. We've also added two lines (median), one vertical and the other horizontal, as reference points.

To understand the chart :

- The further to the right the points are located, the more seats the cantons have.

- The higher up the graph you go, the higher the voter turnout.

Our findings:

- We observe that, in general, the smaller the number of seats for a canton, the greater its participation in the vote.

- Cantons with a high turnout rate, i.e. above the median, such as Schaffhausen, Obwalden, etc., don't have many seats in parliament; they are situated to the left of the median and therefore their number of seats is below average.

- Conversely, cantons with a high number of seats, such as Zürich, which has over 35 seats in parliament, have a relatively low voter turnout. Here, Zürich is well to the right of the median, indicating that it has many seats, but below the median, so votes less than average.

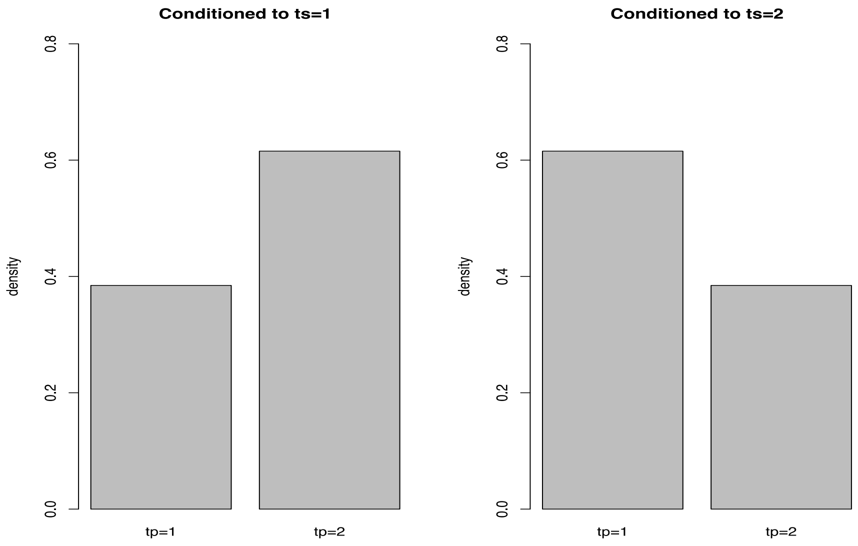


Figure 9 : conditioned to ts=1&2 graphs.

We can see the two conditional density graphs. R formulas for median calculations have enabled us to define what is considered ts=1 (few seats) and ts=2 (many seats), as well as tp=1 (moderate participation) and tp=2 (high participation).

When ts=1, i.e. when cantons have few seats, tp=2 has a higher density than tp=1, showing that when cantons have few seats, their turnout is high (tp=2).

Conversely, in the following graph, when ts=2 (cantons with many seats), then tp=1 has a higher density than tp=2, meaning that the voting participation of cantons with many seats in parliament is moderate (tp=1).

Conclusion:

This analysis enabled us to contradict our basic hypothesis, and thus to observe that cantons with many seats in parliament have a less pronounced turnout in votes than cantons with few seats.

4.3. Swiss (richesse cantional PIB)

We set out with a very general question being:

Is there a correlation between wealth and voting percentage?

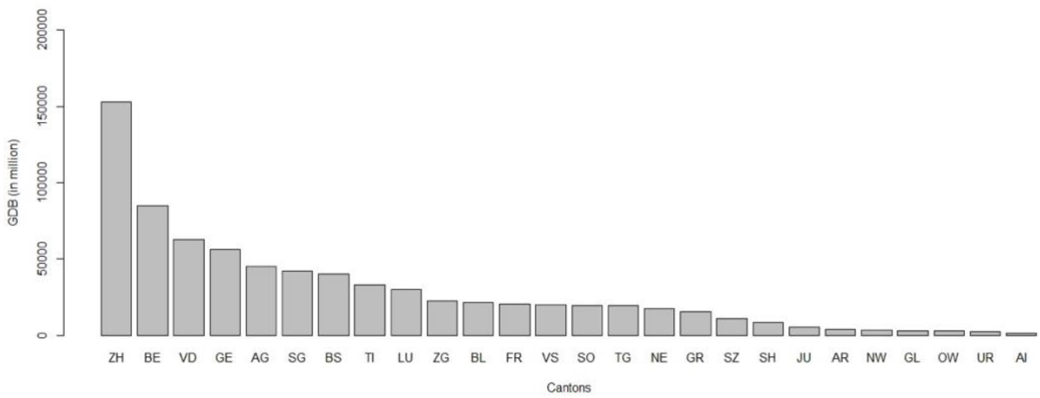


Figure 10: Canton's GDB

We decided to begin by checking if there was a difference between the cantons' wealth by comparing their GDB .We retrieved the data from the administration's website, which enabled us to draw the following bar chart:

There are several things to note:

- The wealth gap between the two richest cantons is quite important.
- The gap narrows over time
- The gap between the most and least affluent cantons is enormous.

Let's take a closer look at the data: (millions of francs):

The calculations were made on the R software except for the gap at the end:

- The median is 19,768: half the cantons have a GDP greater than or equal to this figure, and the other half less than or equal to it.

- The mean is 28,589: so half the cantons (at least) have a lower GDP.

- The highest GDP is 152,547 in Zurich, and the lowest is 1,155 in Appenzell Innerrhoden. The difference between the richest and poorest cantons is 151,392 (it could perhaps be said that the GDP of the canton of AI is negligible compared with Zurich).

We can therefore see that there is a real inequality between the wealth of the cantons. We can now move on to the analysis of vote distributionr. To do this, we've also retrieved data from the admin site and plotted a histogram on R to highlight the frequency of voting percentages.

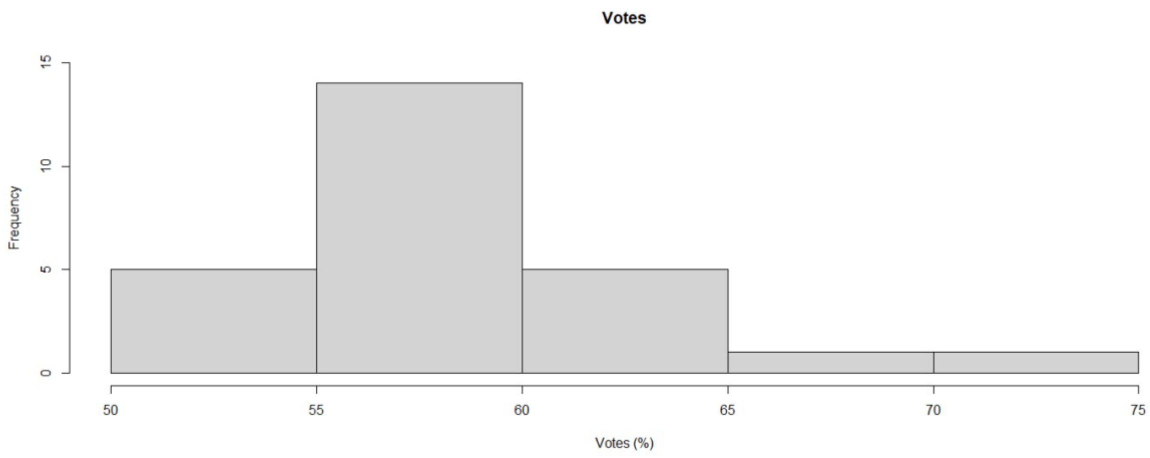


Figure 11: Votes frequency distribution

We note that :

- The highest frequency of vote is located between [55 ;60] % , visually located around 13. We can conclude that in half of the cantons (at least), between 55 and 60% of the population votes.

- The percentages are exclusively between 50 and 75%, so no canton has a voting percentage higher than 75 or lower than 50.

- In addition, there is a shallow frequency for the 65-70% and the 70-75% bracket (probably 1, to be verified).

We have also calculated the average voting percentage for Switzerland as a country, which was around 59%, which is well within the range we thought.

After analyzing the two variables on their own, we'll now link them to verify our hypotheses. To do that, we began by creating a bar chart correlating them.

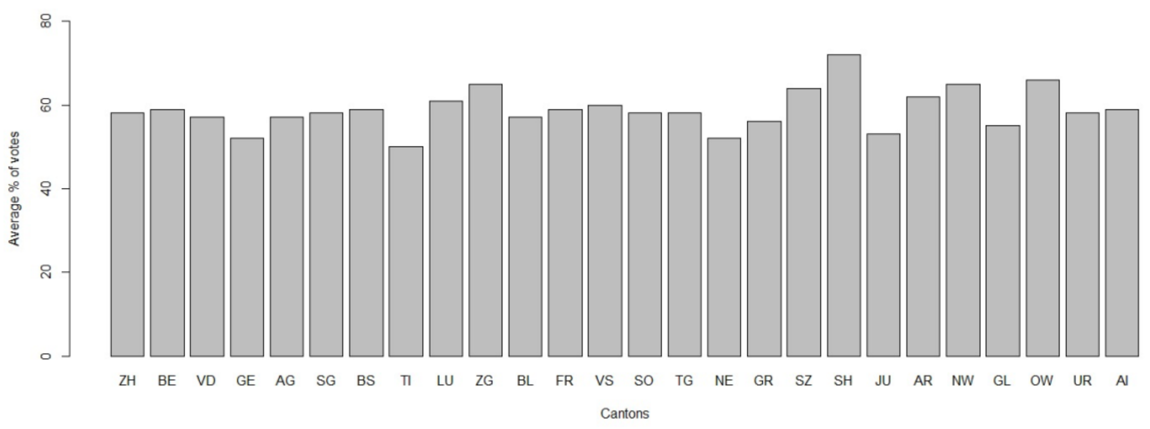


Figure 12: Canton's average vot percentage

We can notice a few points:

- We can see that there are no huge variations between cantons. However, we notice a peak in the canton of Schaffhausen with more than 72% of votes (percentage found thanks to a line of code). This canton's GDP is 8250(millions) and it is the 19th richest canton out of the 26.

- Using a line of code, we looked at the minimum percentage of votes (roughly): which is 50%. It corresponds to the canton of Ticino, which has a GDP of 33,181 (millions). It's in 8th position in the ranking.

As the graph isn't precise enough, we created a scatterplot with a horizontal line representing the average number of votes (in blue) and a vertical line representing the average GDP (in red).

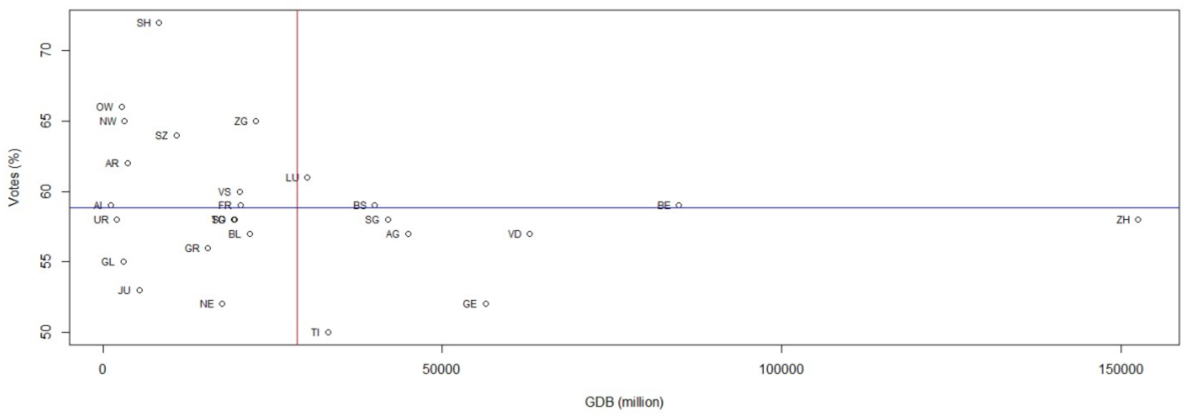


Figure 13: Votes percentage according to the GDB

What we observe:

- When we look at the cantons with an above-average GDP, in terms of voting they are around the average, with two peaks well below the average, but they are still very homogeneous.

- When we look at the cantons with below-average GDP, we can see that they are very heterogeneous, which leads us to look at the graph the other way round: instead of looking at average GDP, we'll look at average votes.

We can see that the most of cantons above the voting average are mostly below the GDP average.

In conclusion, we note that the canton that votes the least is above the GDP average (TI) and that the canton that votes the most is below (SH). This first approach could lead us to hypothesize that the poorest cantons vote more. However, if we look at the extremes, the richest and poorest cantons are within the voting average. Looking more generally, the richest cantons were on average around the voting average, while the poorest cantons were widely scattered between the average and the two other extremes. This leads us to conclude that there is no real link between voting and GDP.

4.4. Limitations

Our work has encountered a number of limitations that deserve to be addressed, since it would be unwise to claim that our analysis has enabled us to draw a complete social and economic impact on voter turnout in our population.

Firstly, we have chosen to focus on three variables and relate them to the percentage of voter turnout in each canton. However, there are many other relevant variables that we have not taken into account in this analysis, such as level of education or incentives to vote in each canton.

Secondly, we chose to focus on Switzerland to obtain reliable data, but there is no guarantee that the population of other countries would act in the same way as the Swiss population. Similarly, our choice of individuals was cantons. However, the voting performance of each canton does not allow us to observe the voting involvement of each individual, which could appear unfair since an individual who is very involved in voting will be hidden in the mass of actions of his canton.

Thirdly, although Switzerland is a quadrilingual country, we did not take Romansh into account in the analysis of the language variable. This is because Romansh is only spoken in Graubünden, and less than 1% of the Swiss population speaks it, making it of little relevance to distinguish this language. What's more, Graubünden is a trilingual canton, which also makes it difficult to obtain data specific to Romansh.

Finally, our project involved the use of the statistical method and not other methods of analysis, which might give a different perspective on things.

5. Discussion

This project enabled us to highlight and analyze the impact of various economic and social variables on the exercise of democratic power in Switzerland, using the 26 cantons that make up our country as individuals.

Firstly, we looked at the link between cantonal language and voter turnout. The results confirm our hypothesis: Swiss-German cantons, especially those with small populations, have higher voter turnout than French- and Italian-speaking cantons.

Secondly, we wanted to analyze the relationship between the total number of seats each canton has in parliament and its participation in votes. The main observation contradicts our basic hypothesis: the more seats a canton has in parliament, the less votes it participates in. Conversely, the fewer seats a canton has in parliament, the higher the percentage of voter turnout.

Thirdly, we investigated whether the wealth (in our report, GDP is used as an indicator) of each canton has an influence on voter turnout in those same cantons. What we found was that the lower a canton's GDP, the higher its voter turnout, which contradicts our initial hypothesis.

What's more, thanks to the three Scatterplot graphs, we can see that for each of our X variables, the points corresponding to each canton keep the same ordinates (since our Y variable, i.e. the percentage of voter turnout for each canton, is the same) but have a different abscissa. This allows us to assert that each of our X variables (Languages, Seats or GDP) has a different influence on our Y variable (voter turnout), since if this were not the case, the point corresponding to a canton would have the same ordinate and the same abscissa in each of our three Scatterplots. However, it is important to note that the difference in the impact of each of our three variables on voter turnout is relatively subtle.

In conclusion, the question of the use of democratic power in Switzerland is a pertinent one, and one that continues to be exploited by the FSO and other institutions, since it lies at the very root of our country's economic and structural future.

6. Bibliography

https://www.bfs.admin.ch/bfs/fr/home/statistiques/politique/votations/participation.assetdetail.23909927.html

https://www.bfs.admin.ch/bfs/fr/home/statistiques/economie-nationale/comptes-nationaux/produit-interieur-brut-canton.assetdetail.28405427.html

https://www.admin.ch/gov/fr/accueil/documentation/communiques.msg-id-84941.html

https://www.bfs.admin.ch/bfs/fr/home/statistiques/politique/votations/participation.assetdetail.27925359.html

A. Appendix –R Code

```
getwd("/Users/albamileci")

setwd("/Users/albamileci/Desktop/UNINE/Statistique:Info/Statistique descriptive/projet")

#Langues des cantons:

install.packages("readxl")

library("readxl")

Votation= read_excel("votations_4.xlsx")

#R1

n=dim(Votation)

ncol(Votation)

nrow(Votation)

colnames(Votation)

# min et max pop

min(Votation$pop_inscrite)

max(Votation$pop_inscrite)

which.max(table(Votation$pop_inscrite))

which.min(table(Votation$pop_inscrite))

#min et max moyenne

min(Votation$moyenne)

max(Votation$moyenne)

which.max(table(Votation$moyenne))

which.min(table(Votation$moyenne))

#graphique géneral

x1=Votation$pop_inscrite

y1=Votation$moyenne

plot(x1, y1, col = "red", pch = 16, main = "Registered population vs average canton participation",
xlab = "Registered population", ylab = "Average participation")

#Aggiungere etichette per i cantoni

text(Votation$pop_inscrite, Votation$moyenne, labels = Votation$canton, pos = 1, col = "black")

#Graphique

#1)Variable x2 = suisse allemanique

x2= Votation$pop_inscrite[1:20]

y2= Votation$moyenne[1:20]
```

```
#2)Variable x3 = suisse française

x3=Votation$pop_inscrite[21:25]

y3=Votation$moyenne[21:25]

#3)Variable x4 = suisse italienne

x4=Votation$pop_inscrite[26:26]

y4=Votation$moyenne[26:26]

#4)Données noms cantons

canton_Allemand_x2 = Votation$canton[1:20]

canton_Romand_x3 = Votation$canton[21:25]

canton_Italien_x4 = Votation$canton[26:26]

#5)plot

plot(x2, y2, col = "blue", pch = 16, xlim = c(min(c(x2, x3, x4)), max(c(x2, x3, x4))), ylim = c(min(c(y2, y3, y4)), max(c(y2, y3, y4))), xlab = " Registered population", ylab = "Average participation %", main = "Graphique Cantons: German, French and Italian")

points(x3, y3, col = "red", pch = 16)

points(x4, y4, col = "green", pch = 16)

abline(v= mean(x1), col= "pink")

abline (h=mean(y1), col="pink")

#6)Canton

text(x = x2, y = y2, labels = canton_Allemand_x2, pos = 1, col = "blue")

text(x = x3, y = y3, labels = canton_Romand_x3, pos = 3, col = "red")

text(x = x4, y = y4, labels = canton_Italien_x4, pos = 4, col = "green")

#7)Legend

legend("topright", legend = c("German", "French", "Italian"), col = c("blue", "red", "green"), pch = 16)

#graphique séparé

par(mfrow=c(3,1))

plot(x2, y2, col = "blue", pch = 16, xlim = c(min(c(x2, x3, x4)), max(c(x2, x3, x4))), ylim = c(min(c(y2, y3, y4)), max(c(y2, y3, y4))), xlab = "Registered population ", ylab = "Average participation %", main = "German Switzerland")

plot(x3, y3, col = "red", pch = 16, xlim = c(min(c(x2, x3, x4)), max(c(x2, x3, x4))), ylim = c(min(c(y2, y3, y4)), max(c(y2, y3, y4))), xlab = "Registered population ", ylab = "Average participation %", main = "French Switzerland")

plot(x4, y4, col = "green", pch = 16, xlim = c(min(c(x2, x3, x4)), max(c(x2, x3, x4))), ylim = c(min(c(y2, y3, y4)), max(c(y2, y3, y4))), xlab = "Registered population ", ylab = "Average participation %", main = "Italian Switzerland")
```

```
#Histogram

#histogram pop_inscrite + canton

barplot(height = Votation$pop_inscrite, names.arg = Votation$canton, col = "lightgreen", main =
"Population registered by canton", xlab = "Canton", ylab = "Registered population")

#histogram moyenne +canton

barplot(height = Votation$moyenne, names.arg = Votation$canton, col = "lightblue", main =
"Average participation by canton", xlab = "Canton", ylab = "Average participation %")

#1) histogrem moyenne + canton

barplot(height = Votation$moyenne, names.arg = Votation$canton, col = "lightblue", main =
"Average participation by canton", xlab = "Canton", ylab = "Average %")

# 2) deuxième partie avec 'pop_inscrite'

par(new = TRUE)

barplot(height = Votation$pop_inscrite, col = "lightgreen", axes = FALSE)

# 3) axe pour deuxième partie

axis(side = 4)

mtext("Pop_Inscrite", side = 4, line = 3)

#Tableau de contengence

#1)caractére pop_inscrite

contingency_pop_inscrite=table(Votation$pop_inscrite)

print(contingency_pop_inscrite)

#2) caractère moyenne

contingency_moyenne = table(Votation$moyenne)

print(contingency_moyenne)

#3) densité pop_enscrit

densité_pop_inscrite = density(Votation$pop_inscrite)

#4) densité moyenne

densité_moyenne = density(Votation$moyenne)

# Frequenze cumulate

freq_cumulative_pop = cumsum(contingency_pop_inscrite)

freq_cumulative_moyenne=cumsum(contingency_moyenne)

# Pourcentage

pourcentage = prop.table(contingency_pop_inscrite) * 100

tableau_complet = cbind(contingency_pop_inscrite, pourcentage)

colnames(tableau_complet) = c("Contingency_pop_enscrit", "Pourcentage")
```

```
tableau_complet2= cbind(contingency_moyenne, pourcentage)
colnames(tableau_complet) = c("Contingency_moyenne", "Pourcentage")
# tableau complet
print(tableau_complet)
prop.table(tableau_complet)
print(tableau_complet2)
prop.table(tableau_complet2)
#R3
#1)Variable moyenne
F_moyenne= ecdf(Votation$moyenne)
par (mfrow= c(1,1))
plot(F_moyenne, xlab="Votation$moyenne", ylab="Votation$moyenne", main="Cdf")
f_moyenne = prop.table(table(Votation$moyenne))
par(mfrow = c(2, 1))
barplot(f_moyenne, xlab = "Moyenne_group", ylab = "Probabilities", main = "Density")
plot(F_moyenne, xlab = "Moyenne_group", ylab = "Probabilities", main = "Cdf")
quantile(F_moyenne, prop = 0.5)
abline(h = 0.5, lty = 2, lwd = 2, col = "red")
abline(v = c(0.5000, 0.5700, 0.5800, 0.6075,0.7200), lty = 2, lwd = 2, col = "violet")
#hist
par(mfrow = c(1, 2))
hist(Votation$moyenne)
barplot(f_moyenne, xlab = "Moyenne", ylab = "Probabilities", main = "Density")
par(mfrow = c(1, 2))
hist(Votation$pop_inscrite)
barplot(f_pop, xlab = "Population", ylab = "Probabilities", main = "Density")
#densité jointe
#22) Construire la densité jointe pour les variables X et Y.
X1=x1
Y1=y1
Joint_density=table(X1,Y1)
print(Joint_density)
TP=Votation$pop_inscrite
```

```
TM=Votation$moyenne

Cant=Votation$canton

aa=table(Votation$pop_inscrite,Votation$moyenne)

TM_med=median(TM)

TM_med

tm=TM

tm[TM<TM_med]=1

tm[TM>=TM_med]=2

TP_med=median(TP)

TP_med

tp=TP

tp[TP<TP_med]=1

tp[TP>=TP_med]=2

A=data.frame(cbind(Cant,tp,tm))

Contengency=table(A$tp,A$tm)

Contengency

Density=Contengency/nrow(Votation)

Density

#Faire des densités conditionelles

tm_1=prop.table(table(tp[tm==1]))#montre la densité conditionelle du caractère "population inscrite" quand "Moyenne de participation par contans" est = 1

tm_1

rownames(tm_1)=c ("tp=1","tp=2")

barplot(tm_1, ylim=c(0,0.8), main ="Conditioned to tm = 1", ylab="Density")

tm_2=prop.table(table(tp[tm==2]))#montre la densité conditionelle du caractère "population inscrite" quand "Moyenne de participation par contans" est = 1

tm_2

rownames(tm_2)=c ("tp=1","tp=2")

barplot(tm_2, ylim=c(0,0.8), main ="Conditioned to tm = 2", ylab="Density")

par(mfrow= c(1,2))

barplot(tm_1, ylim=c(0,0.8), main ="Conditioned to tm = 1", ylab="Density")

barplot(tm_2, ylim=c(0,0.8), main ="Conditioned to tm = 2", ylab="Density")
```

#Sièges au Parlement des cantons:

```
install.packages("readxl")

library(readxl)

Base=read_excel("Sièges_cantons.xlsx")

#2) Définir nos variables

Cantons=Base$Cantons

Cantons

Average_voting=Base$`Moyennes participation votes`

Average_voting

Voting_figures=Base$Chiffres

Voting_figures

Total_seats=Base$`Total sièges`

Total_seats

#3) Extraire les noms de notre tableau

colnames(Base)

#4) Extraire le nombre le plus petit des sièges détenus par les cantons

min(Total_seats)

#5)Analyser les 10 premières et dernières observations du tableau.

n=dim(Base)[1]

n

Base[1:10,]

Base[(n - 9):n, ]

#6) Voir la variabilité du caractère "Total_sièges_canton" avec un tableau de contingence.

contingency_total_seats = table(Total_seats)

contingency_total_seats

#7)Faire un tableau de fréquence pour cette même variable.

Frequency_total_seats=contingency_total_seats/n

Frequency_total_seats

#8)Visualiser les tableaux de contingence et de fréquence en même temps

par(mfrow = c(1, 2))

barplot(contingency_total_seats, ylab = "number of cantons by number of seats", xlab = "number of seats",main = "Contingency")

barplot(Frequency_total_seats, ylab = "density", xlab = "number of seats", main = "Frequency")

#9)Extraire le canton qui a le plus de sièges ainsi que le nombre de sièges possédé.
```

Cantons[which.max(Total_seats)]

max(Total_seats)

#10)La proportion de ce maximum de sièges parmis les sièges totaux.

max(table(Cantons))/nrow(Base)

#11) Extraire le minimum de sièges ainsi que les cantons ayant ce nombre de sièges.

Min_number_seats=min(Total_seats)

Cantons[Total_seats==Min_number_seats]

#12) La proportion de ce minimum de sièges parmis les sièges totaux.

Min_number_seats/nrow(Base)

#13)Fonction de densité pour la variable "sièges"

Density_function_seats=ecdf(Total_seats)

par(mfrow = c(1, 1))

plot(Density_function_seats, xlab = "Number of seats", ylab = "Probabilities", main = "Cumulative distribution function")

#Visualiser la médiane du graphique.

quantile(Density_function_seats, prob=0.5)

abline(h = 0.5, lty = 2, lwd = 2, col = "red")

abline(v=7.5, lty = 2, lwd = 2, col = "purple")

#17)Histogramme pour la variable "Nombre de sièges" comparé à la densité

par(mfrow = c(1, 2))

hist(Total_seats,xlab="Number of seats", main="Histogram")

barplot(Density_function_seats, xlab = "Number of seats", ylab = "Probabilities",main = "Density")

#18) Probabilité de tomber sur un canton avec au moins 4 sièges.

Density_function_seats(4)

#19) Probabilité qu'en choisissant 1 canton, on tombe sur un canton avec un nombre de sièges entre 2 et 10.

Density_function_seats(10)-Density_function_seats(2)

#20)Combien de sièges au moins ont le 30% des cantons avec le plus de sièges?

quantile(Density_function_seats, prob=0.7)

#21) Création d'un scatterplot réunissant deux variables dans un graphique et montrer les moyennes des variables

X = Total_seats

Y = Average_voting

plot(X,Y, main="Scatterplot X&Y",xlab ="X=number of seats", ylab= "Y=voting participation")

```
text(X, Y, labels = Cantons, pos = 1, cex = 0.6, col = "blue")

abline(v = mean(X), col = "red")

abline(h = mean(Y), col = "red")

#22) Construire la densité jointe pour les variables X et Y.

TS=Base$`Total sièges`

TP=Base$`Moyennes participation votes`

Cant=Base$Cantons

aa=table(TS,TP)

TP_med=median(TP)

#On fait la médiane pour savoir ce qui va être considéré comme une participation modérée (< médiane)et une participation élevée (> médiane).

TP_med

tp=TP

tp[TP<TP_med]=1 #On définit ce qui est considéré comme participation modérée =1

tp[TP>=TP_med]=2 #On définit ce qui est considéré comme participation élevée =2

TS_med=median(TS)#On fait la médiane pour savoir ce qui est considéré comme peu de sièges (=1) et beaucoup de sièges (=2).

TS_med

ts=TS

ts[TS<TS_med]=1 #peu de sièges -> 1

ts[TS>=TS_med]=2 #beaucoup de sièges -> 2

A=data.frame(cbind(Cant,ts,tp)) #Tableau de contingence joint

Contingency=table(A$ts,A$tp)

Density= Contingency/nrow(Base) #Tableau de densité jointe

#23) Faire des densités conditionnelles

tss_1=prop.table(table(tp[ts==1])) #Montre la densité conditionnelle du caractère "Moyenne participation votes" quand la variable "Total sièges" est =1 (peu de sièges).

tss_1

rownames(tss_1) = c("tp=1", "tp=2")

barplot(tss_1, ylim = c(0, 0.8), main = "Conditioned to ts=1", ylab = "density") #On fait un graphique qui représente cette densité conditionnelle.

tss_2=prop.table(table(tp[ts==2]))#Montre la densité conditionnelle du caractère "Moyenne participation votes" quand la variable "Total sièges" est =2 (beaucoup de sièges).

tss_2

rownames(tss_2) = c("tp=1", "tp=2")
```

```
barplot(tss_2, ylim = c(0, 0.8), main = "Conditioned to ts=2", ylab = "density")
#On fait un graphique qui représente cette densité conditionnelle.
par(mfrow = c(1, 2))
barplot(tss_1, ylim = c(0, 0.8), main = "Conditioned to ts=1", ylab = "density")
barplot(tss_2, ylim = c(0, 0.8), main = "Conditioned to ts=2", ylab = "density")
```

#PIB cantons:

```
install.packages("readxl")
library(readxl)
PIB_Votes = read_excel("PIB_Votations.xls")
## Creation des variables##
Cantons = PIB_Votes$Cantons
PIB = PIB_Votes$`PIB(millions de francs)`
Votes = PIB_Votes$Votations
## Analyse de la variable PIB ##
Mediane_PIB = quantile(PIB,0.5)
# donc 50% des cantons ont un PIB <= 19 768  million et 50% >= 19 768  million
Moyenne_PIB = mean(PIB)
Moyenne_PIB # =~ 28 589.62 millions
Total_PIB = sum(PIB)
Total_PIB # = 743 330 millions
Max_PIB = max(PIB)
Max_PIB # = 152 547 millions
Min_PIB = min(PIB)
Min_PIB # = 1 155 millions
## Analyse de la variable Votation ##
Mediane_Votation = quantile(Votes,0.5)
# 50% des cantons ont un % de vote <= 57.3% et 50% ont >= 57.3%
Moyenne_Votation = mean(Votes)
Moyenne_Votation # =~ 57.9 %
Max_Votation = max(Votes)
Max_Votation # =~ 72 %
Min_Votation = min(Votes)
Min_Votation # =~ 50.3 %
```

```
## Graphs ##
# graph 1 #
plot(PIB,Votes, xlab = "GDB (million)", ylab = "Votes (%)")
abline(v = Moyenne_PIB, col = "red")
abline(h = Moyenne_Votation, col = "red")
text(PIB,Votes,label=Cantons, pos = 2, cex=0.5) # pos == position
# graph 2 #
par(mfrow=c(1,1))
# graph 2.1 #
barplot(Votes, xlab = "Cantons", ylab = "Average % of votes", ylim = c(0,70), names.arg = Cantons)
# canton en anglais y avait plusieurs trucs.
# sur google translate y avait canton, township, municipality...
# quand je rearrange par ordre les bar, le noms des cantons bouge pas alors j'ai garde comme ca
# graph 2.2 #
barplot(PIB, ylim = c(0,200000), names.arg = Cantons, xlab = "Cantons", ylab = "GDB (in million)")
# GDB (Gross Domestic Product) == PIB en anglais
# graph 3 #
hist(Votes, main = "Votes", ylim = c(0,15))
# graph 4 #
Flog=ecdf(log10(PIB)) # utiliser ecdf() pour construire les fonctions de repartitions
Flog
par(mfrow=c(1,1))
plot(Flog)
abline(h = 0.5, col = "red")
abline(v = 4.0, col = "red")
hist(Flog, ylim = c(0,15),main="GDB's log10", xlab = "log10(GDB)")
```

Déclaration d'honneur

Par la présente, j'affirme avoir pris connaissance des documents d'information et de prévention du plagiat émis par l'Université de Neuchâtel et m'être renseigné-e correctement sur les techniques de citation. J'atteste par ailleurs que le travail rendu est le fruit de ma réflexion personnelle et a été rédigé de manière autonome.

Je certifie que toute formulation, idée, recherche, raisonnement, analyse ou autre création empruntée à un tiers est correctement et consciencieusement mentionnée comme telle, de manière claire et transparente, de sorte que la source en soit immédiatement reconnaissable, dans le respect des droits d'auteur et des techniques de citations.

Je suis conscient-e que le fait de ne pas citer une source ou de ne pas la citer clairement, correctement et complètement est constitutif de plagiat. Je prends note que le plagiat est considéré comme une faute grave au sein de l'Université. J'ai pris connaissance des risques de sanctions administratives et disciplinaires encourues en cas de plagiat (pouvant aller jusqu'au renvoi de l'université).

Je suis informé-e qu'en cas de plagiat, le dossier sera automatiquement transmis au rectorat. Au vu de ce qui précède, je déclare sur l'honneur ne pas avoir eu recours au plagiat ou à toute autre forme de fraude.

Neuchâtel, 20 Décembre 2023

Alba Mileci, Claire Petrini,
Laila Ibrahim, Clara Comanescu,
Lisa Zaepfel et Mélissa Sönmez.

Texte basé sur le document « Déclaration sur l'honneur » du Rectorat de l'Université de Neuchâtel.