

Youth worldview analysis & insights: have we changed from 2002 to 2019?

Yaryna Lyba, Anastasiia Livochka

Abstract

We live in a really young country that is built every day. People that were 17-25 in 2002 are making now decisions that influence our lives. 16-22 years old people are on the way of developing themselves as citizens with their principles. If we could understand the young generation now, we could help them to go in the right way and prevent all the nation from big mistakes that are made again and again throughout the whole of our history.

I. DATA COLLECTION

For our research, we used two data sources. First one dataset we got from the National bank of researches. There we had 519 answers on more than 100 questions from 62 settlements of Ukraine in all its districts. They were collected in the year 2002. Reading the questions, we got interested in what would people response to the same now. We have cast away irrelevant for today questions like about frequency of using the internet. Out of those questions that left we selected 30 best and based on them we constructed our survey. We have sent it to people from different cities and villages. In two weeks we managed to get 500 answers from the whole Ukraine (except Crimea).

II. DATA PREPARATION

Data from the second dataset was biased; we got more answers from people from the west; there were many answers from UCU students. So we worked a lot on normalization. The first thing that we checked was the distribution of male-female interviewees. It appeared to be almost perfect.

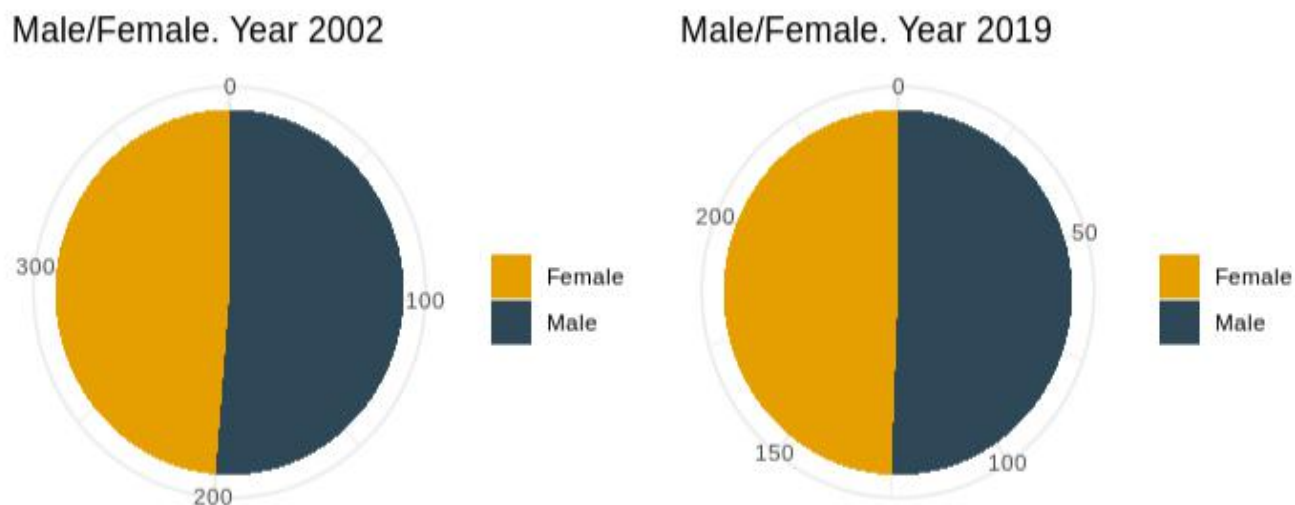


Figure 1: Distributions of answers from men and women 2002 vs 2019

But the most challenging thing for us was to ensure that we have no bias in the distribution of people by the territory they live. So we divided all districts of Ukraine into four parts: west, east, center, and south. We have calculated the number of people that live in those parts and made the same fractions in our datasets.

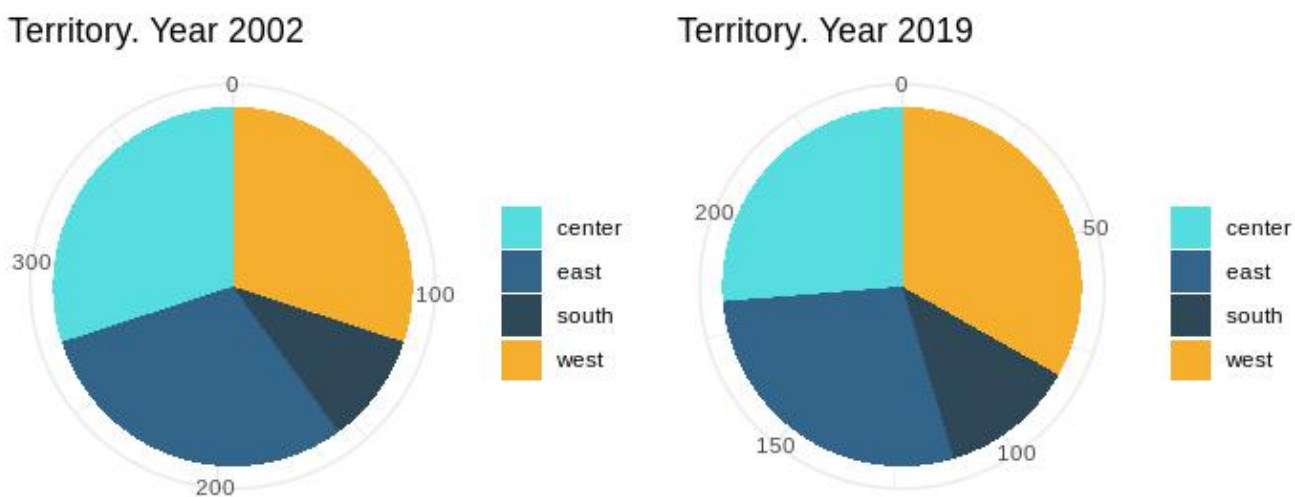


Figure 2: Fractions of people living in different parts of Ukraine in two datasets.

We validate that sample distribution is unbiased by matching election results with answers to the question about next president we have put in our survey. Without normalization we had biased distribution (57,4 percent for Petro Poroshenko and 42,6 for Volodymyr Zelenskyy), but after data preprocessing we got results close to real.

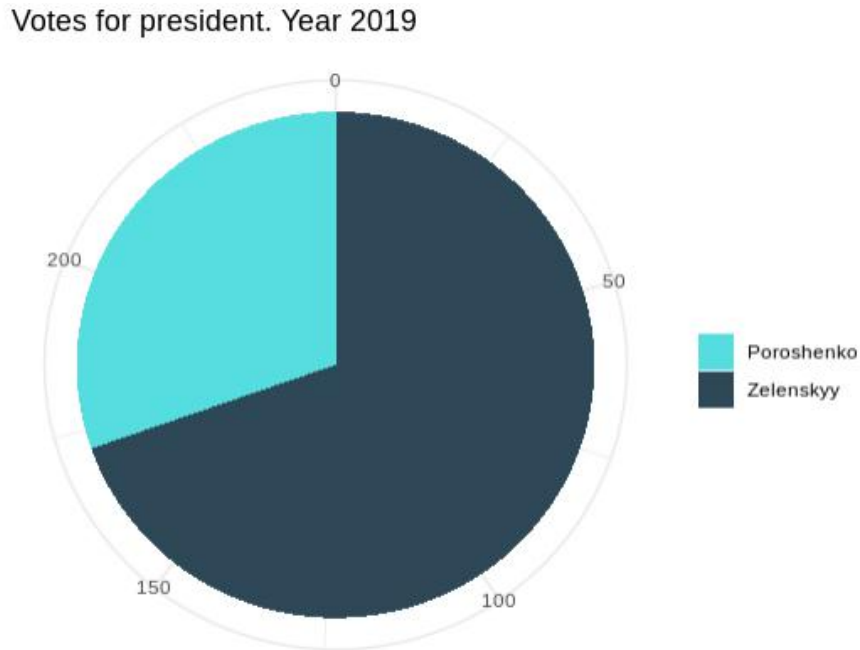


Figure 3: Answers to the question about next president.

III. HYPOTHESIS TESTING: DO WE TRUST LESS?

In both datasets we had questions about level of trust to different institutions. We test whether the average Ukrainians trust levels in 2019 are the same as in 2002 at significance level $\alpha = 0.05$.

For hypothesis testing we use Mann-Whitney-Wilcoxon Test. Using the Mann-Whitney-Wilcoxon Test, we can decide whether the population distributions are identical without assuming them to follow the normal distribution.

Our first test is about army trust levels. $H_0 : \mu_{2019} = \mu_{2002}$, $H_1 : \mu_{2019} > \mu_{2002}$, σ_1^2, σ_2^2 unknown. Wilcoxon rank sum test with continuity correction gave p-value = 0.3691, that is greater than alpha, so we fail to reject null hypothesis. The average army trust level in 2019 is the same as in 2002 at significance level $\alpha = 0.05$.

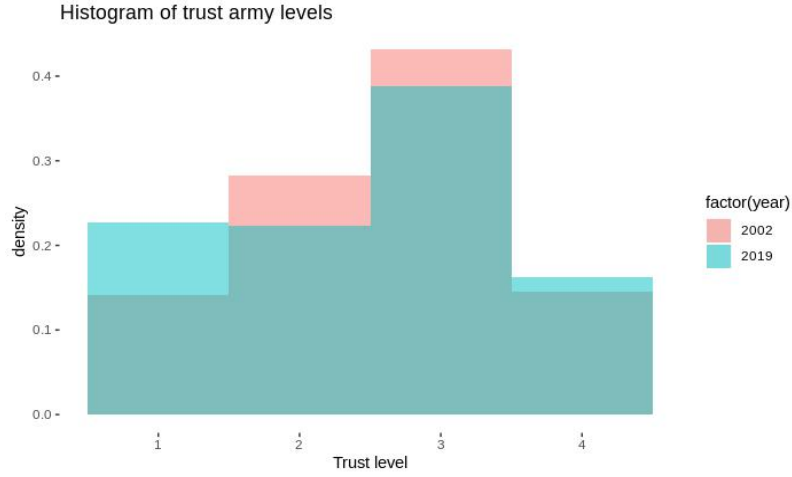


Figure 4: Army trust levels.

Next test is about trust level to president. $H_0 : \mu_{2019} = \mu_{2002}$, $H_1 : \mu_{2019} > \mu_{2002}$, σ_1^2, σ_2^2 unknown. Wilcoxon rank sum test with continuity correction gave p-value = 0.3174, that is greater than alpha, so we fail to reject null hypothesis. The average trust level to president in 2019 is the same as in 2002 at significance level $\alpha = 0.05$. That can be seen on the histogram.

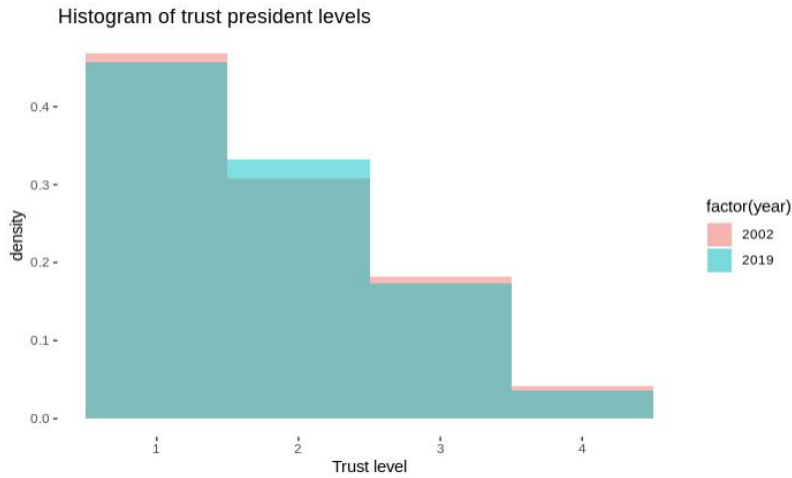


Figure 5: President trust levels.

We also test ZMI trust levels. $H_0 : \mu_{2019} = \mu_{2002}$, $H_1 : \mu_{2019} > \mu_{2002}$, σ_1^2, σ_2^2 unknown. Wilcoxon rank sum test with continuity correction gave p-value $< 2.2e-16$, that is smaller than alpha, so we reject null hypothesis. The average ZMI trust level at 2019 is lower than in 2002. That can be seen on the histogram.

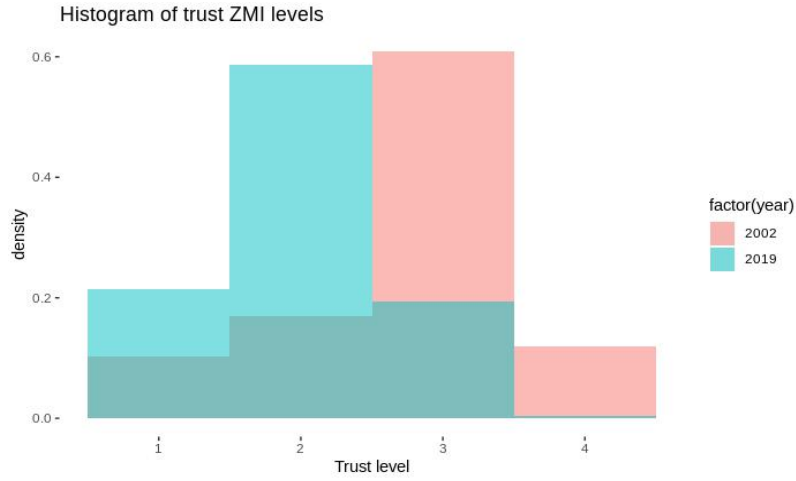


Figure 6: ZMI trust levels.

One more interesting test is about life satisfaction levels. $H_0 : \mu_{2019} = \mu_{2002}$, $H_1 : \mu_{2019} < \mu_{2002}$, σ_1^2, σ_2^2 unknown. Wilcoxon rank sum test with continuity correction gave p-value = 2.545e-05, that is smaller than alpha, so we reject null hypothesis. The average life satisfaction level at 2019 is greater than in 2002. That can be seen on the histogram.

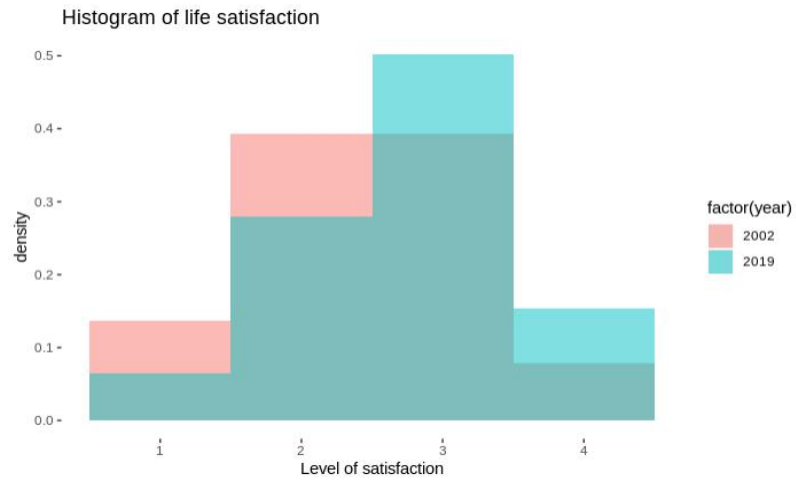


Figure 7: Life satisfaction levels.

Here are some other interesting changes. Corruption occurrence is much less now. The level of church trust decreased also very much.

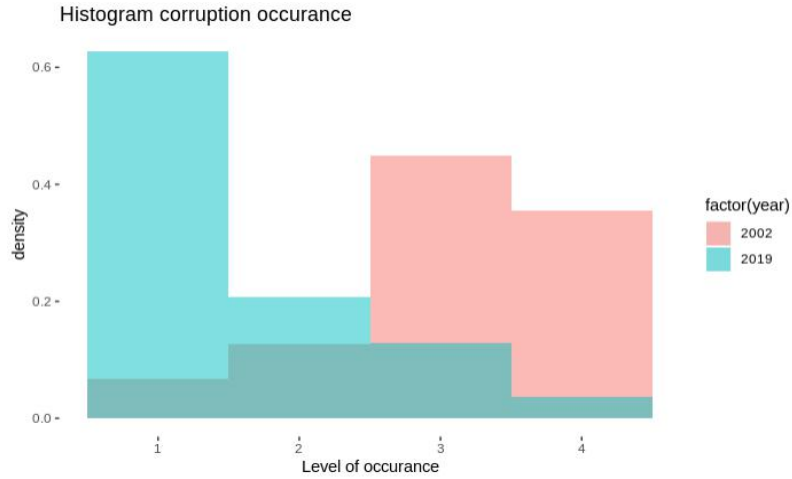


Figure 8: Corruption occurrence levels.

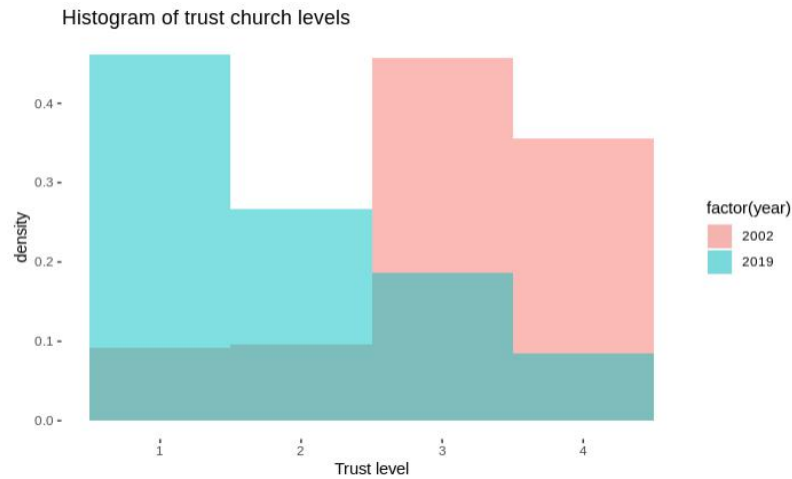


Figure 9: Church trust levels.

IV. POLICY ANALYSIS WITH POOLED CROSS SECTION: WHETHER THE WAR IN THE EAST INFLUENCED EASTERN PEOPLE'S PERCEPTION OF RUSSIA?

Eastern region of Ukraine always had some noticeable cultural differences. It is located close to the Russian Federation, former Empire with appropriate politics, and has an especially complex history. It fell under Russian imperial rule by the late 17th century, much earlier than west of Ukraine. The effect of Russian propaganda, widespread settlement of the East by Russians and other assimilation policies lasts even after the fall of Soviet Union. People here have generally supported more pro-Russian politicians and somehow tend to

Russia, while the west of Ukraine mostly had anti-Russian sentiments historically.

Ukrainian-Russian relations had periods of ties and tensions after the Soviet Union's dissolution but generally were cooperative because of inherited joint facilities of production. For the people from the east, it might even seem friendly. Russia disclosed its real face at Euromaidan; Ukraine lost control of Crimea and part of Donbass because of Russian aggression. We consider the war as the most influence event of the recent years.

A. The effect of informational war on eastern people perception of Russia

Let us firstly look at how youth generally changed its opinion about Ukraine and Russia:

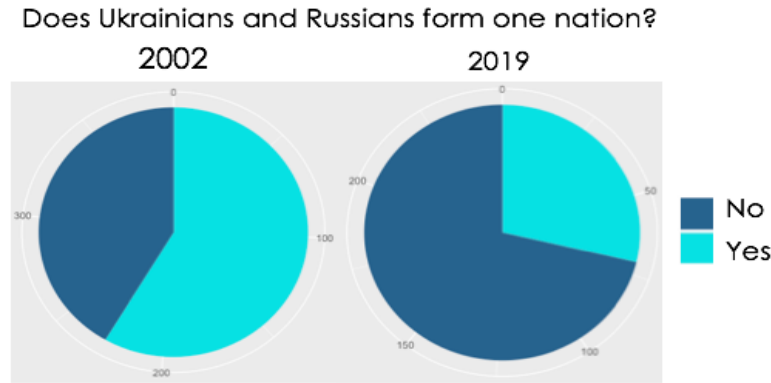


Figure 10: Perception of Russia 2002 vs 2019

So, we can conclude that we are moving in a direction to real independence from Russia even in peoples minds.

But what is going on on the east? We see that the east still vote for pro-Russian politicians, we have a lot of aggressive Russian propaganda here and the war. While we have frozen military conflict, let's look on the situation on the informational field. Our assumptions are: west and eastern regions have a parallel trend in Russian sentiments, but, of course, different intercepts; the war mainly affected east perception of Russia. We decided to use dif-in-dif approach for our purposes and build the logit model targeting on variable of apprehension Ukrainians and Russian as the same public, the control group was the sample from central region, the treatment group was the west sample. The model looks like:

$$r_u_same = \beta_0 + \delta_0 y19 + \delta_1 east + \delta_2 y19 * east + \dots + u$$

where δ_2 has become known as the difference-in-differences estimator and it measures the changes in eastern people relation to Russia due the war, including informational. For solving the problem of over-fitting, we included only significant factors from the full model into the final one.

The results looks not optimistic, you can find them in the end of the section. It requires further analysis, but **from our results it seems like Ukraine loses information warfare**. δ_2 positively correlate with target variable and significant at $\alpha = 0.1$, what means that having other factors fixed people from the east in 2019 feel more close to Russia.

B. The effect of war on happiness of the people

Do we really feel the war? Some of the surveyed people in our data are closer to the conflict region, some of them have a greater risk to appear in the conflict zone tomorrow, but all of them are not located on the firing line and see the war mainly on TV. We decided to estimate the effect of the war in the region you are locating on your life satisfaction. We have a similar assumption as in section A and the same treatment and control group. We've built the linear model:

$$life_satisf = \beta_0 + \delta_0 y19 + \delta_1 east + \delta_2 y19 * east + \dots + u$$

The interpretation of obtained coefficients is logical, the more you trust your friends, mass-media, and banks, the happier you are; if you want to move from here, you, probably, do not feel yourself happy. The good news is people in 2019 are happier than in 2002; we have a positive trend. δ_2 is not significant even at $\alpha = 0.1$, so we can say that there is no effect of the war in your region on your life satisfaction compared to the effect on people in other regions. Probably, the war is more about nations, than specific regions.

One more model showed: the most influenced by war is the desire to move forever from Ukraine.

Table I:

	<i>Dependent variable:</i>		
	r_u_same	life_satisf	move_forever
	<i>logistic</i>	<i>OLS</i>	<i>logistic</i>
	(1)	(2)	(3)
y2019	−1.685*** (0.467)	0.446*** (0.133)	−0.054 (0.563)
east	1.285*** (0.286)	−0.201* (0.116)	−0.149 (0.471)
poor	−0.210 (0.283)		−0.111 (0.377)
very_poor	0.051 (0.288)		−0.842** (0.408)
higher_edu	−0.469 (0.291)		0.623* (0.368)
with_parents		−0.106 (0.096)	0.376 (0.355)
trust_friends		0.239*** (0.058)	−0.092 (0.248)
trust_ZMI	0.175 (0.141)	0.147*** (0.057)	−0.272 (0.222)
move_forever	0.618* (0.342)	−0.284** (0.134)	
trust_banks		0.126*** (0.037)	0.075 (0.136)
trust_president			−0.345 (0.214)
trust_family			0.623** (0.299)
trust_edu			−0.073 (0.211)
y2019:east	0.981* (0.557)	0.095 (0.191)	1.639** (0.684)
Constant	−0.818* (0.469)	1.137*** (0.259)	−3.032** (1.249)

V. CLUSTERING DATA AND ANALYZING ELECTIONS OUTCOME

A. Finding causalities

Elections in Ukraine was one of the hottest topics of April 2019 with one of hottest discussions about candidates we've ever had with a wide involvement of young people. We

wondered what was the key factors that influenced election's outcomes.

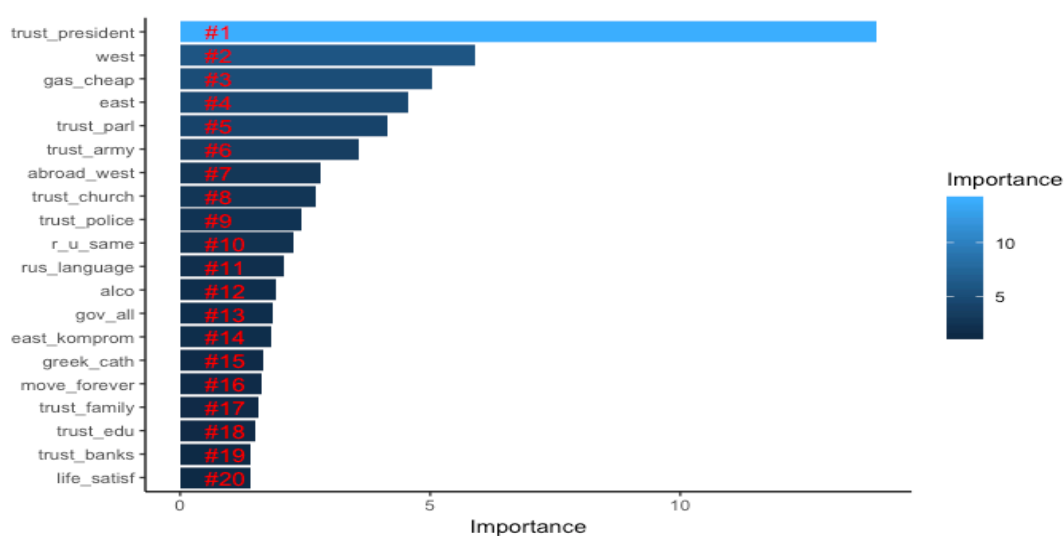


Figure 11: Importance of different factors to elections outcomes

We've built the RandomForest model for prediction of voting and got top-20 important factors.

Accuracy of prediction RF model: 85,3 %

Decision \ Reality	Reality	
	True	False
True	163	9
False	26	49

You can see that voting outcome can be almost pretty well described by the data we have. We decided to run logit model with top-20 variables to understand the causalities and their significance. Look at regression results on the page below.

Interesting that the most significant variable is the trust to the current president what is logical, as he was the opponent to target candidate. Importance of living in west region for voting against target candidate we can explain as the influence of cultural similarity to the current president.

We see the significance and positive correlation of the desire for a state reduction of gas prices and other tariffs with voting for target candidate and it proves the influence of populism in program and ads of target candidate.

Table II:

<i>Dependent variable:</i>	
	factor(zelensk)
trust_president	−2.402*** (0.591)
west	−1.055* (0.606)
gas_cheap	1.368** (0.610)
east	19.317 (1,559.371)
trust_parl	−0.335 (0.541)
trust_army	0.265 (0.351)
abroad_west	−0.657 (0.549)
trust_church	0.236 (0.294)
trust_police	−0.338 (0.388)
r_u_same	1.530* (0.805)
rus_language	0.443 (0.842)
alco	−0.956*** (0.361)
gov_all	0.717 (0.588)
east_komprom	1.730* (0.980)
greek_cath	−0.906 (0.716)
move_forever	2.261** (1.012)
trust_family	0.883** (0.423)
trust_edu	−0.553 (0.413)
trust_banks	0.571 (0.360)
life_satisf	−0.196 (0.379)
Constant	3.911** (1.957)
Observations	247
Log Likelihood	−50.447
Akaike Inf. Crit.	142.894

Note:

*p<0.1; **p<0.05; ***p<0.01

Significance and positive correlation of desire to move forever from Ukraine, find the compromise with separatists and treating Ukrainians and Russians as the same nation with voting for Volodymyr Zelenskiy are partly consequences of his lack of communication with media and public, uncertainty about his view on the situation on Donbass, while Poroshenko had a strong anti-Russian position.

For understanding the significance of alcohol consumption we need further research. We tried to include variables that highly correlate with alcohol in the model, but none of them was significant.

B. Clustering data

We decided to outline the main groups of young people based on their political views to understand the data better. We used voting outcome, level of trust to president and to parliament, opinion about governmental regulation of tariffs, similarity of Russians and Ukrainians, the choice between unpopular, but resultant politics versus just fulfilling the desires of the public and region as input data. We applied k -means algorithm to clustering. Number of clusters k was calculated as the smallest value of k such that the gap statistic is within one standard deviation of the gap at $k + 1$:

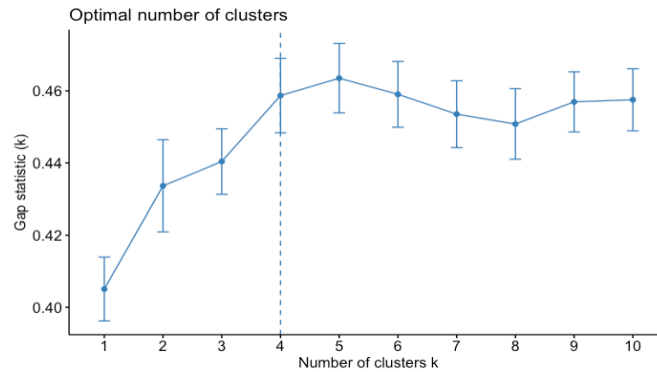


Figure 12: Gap statistic method

We obtained following four clusters with sizes **78, 72, 52, 45** respectively:

The fourth cluster is strongly different from three others; it almost fully consists of Petro Poroshenko electorate. Here people do not consider Russians and Ukrainians one nation, do not want artificial tariffs reduction and, most of them travel to Europe and do not think the government should comply with all requirements of the public.

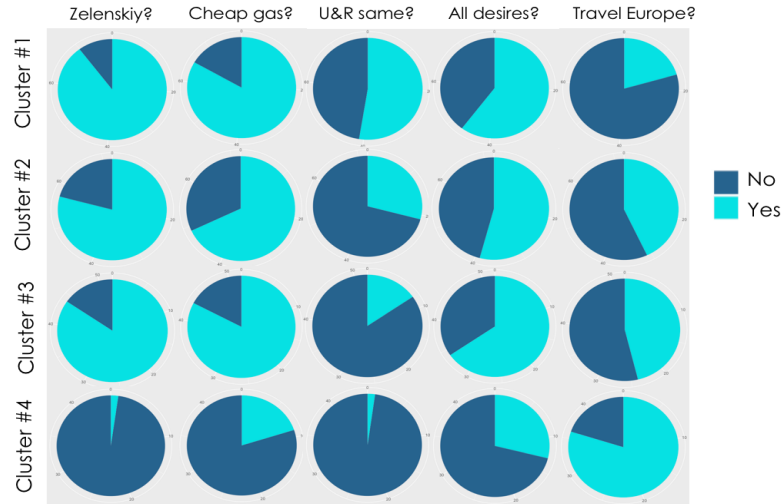


Figure 13: Characterization of clusters

The first three clusters mostly consist of Volodymyr Zelenskiy electorate, and all three clusters support the artificial reduction of tariffs; There is evidence strong negative correlation in the first three clusters of a number of people who travel to Europe with the number that consider Ukrainians and Russians the same. As we see, Zelenskiy electorate consists of three main groups, that are on average much more loyal to Russia, less travel to Europe and are, on average, more vulnerable to populist manipulations.