

Assignment 3

Tori Dykes and Mansi Wadhwa

Friday, November 11, 2016

Introduction

As stated in our last submission, the guiding question for our research project is, **“How does the youth unemployment rate in European countries affect the level and type of political engagement among youths in those same countries?”** To address this questions, we have selected two separate datasets to consider. The first is the European Social Survey, which is a representative survey conducted every two years in up to 36 different countries (the actual number of participating countries tends change slightly between surveys). The survey uses face-to-face interviews to gauge the attitudes, perspectives and beliefs of individuals pertaining to topics like media, politics, social exclusion, and personal values. For our research, we will focus on a selection of the political variables. The second data source we utilize is Eurostat’s data on unemployment in European countries over the last 12 years. The specific measurement we will focus on is the “percentage of active population” that is currently unemployed, specifically for the under-25 grouping. For our purposes, the years 2008 and 2010 will prove particularly relevant, since the economic crisis of 2008 had a significant effect on youth unemployment across Europe.

For the rest of this paper, we will elucidate our data cleaning process, offer some initial descriptive statistics to start to flesh out the topic, and then clarify what we expect our next steps will be for the final paper.

Data Cleaning Process

Since we are focussing on youth unemployment for our analysis, we restricted the data sets to an appropriate age group. The ESS data carries individual-level observations and hence, we are using only the observations falling within the age group of 18 to 25. Similarly, the Eurostat data contains unemployment rates per year per country/geographical division for the total population as well for those less than 25 years of age. We subset our data to contain only the values for those in the group “Less than 25 years of age”.

In order to merge the datasets, both the datasets need to be modified according to our requirements for our research question.

Firstly, the Eurostat dataset currently contains three different measures of youth unemployment. We want to focus only on one measure, namely unemployed youth as a percentage of active population. The data is subset keeping only this single statistic.

Secondly, the time dimension of both datasets needed to be equalised. We restrict the Eurostat dataset to contain only the data for the years for which we have the corresponding ESS survey data. The Eurostat data contains observations from 2006 to 2015 while the ESS was conducted in 7 rounds in the years: 2002, 2004, 2006, 2008, 2010, 2012, 2014. At the same time, ESS data does not contain a year variable but instead, has a variable called ‘essround’ that tells us which survey round and correspondingly, which year a particular observation belongs to. We create a new year variable that takes the information from ‘essround’ and displays the year instead. It is also named the same as the year variable in the second dataset to facilitate merging.

Although our current dataset contains data only until 2012 (owing to the data downloading process), we would like to extract the relevant information from the 2014 ESS dataset and include it in our analysis at the next stage.

Thirdly, different countries/geographical units are covered in the two datasets. While data from Eurostat contains 39 geographical levels, the ESS covers a total of 32 European and non-European countries. We drop all observations pertaining to the European Union as a whole from the Eurostat data as we intend to carry

out our analysis using country-level data, given the diversity of national socio-political systems within the EU. To ensure uniformity and ease of nomenclature, we rename a country (Germany). In a similar vein, we also clean the ESS data by dropping undesired countries, i.e. countries for which we either do not have corresponding data in the other source or countries that do not fall under the ambit of our intended analysis.

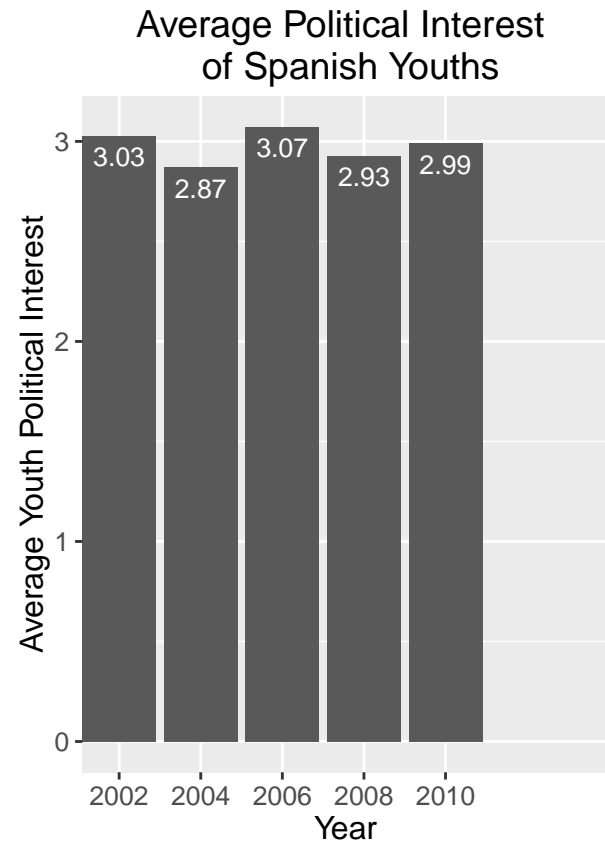
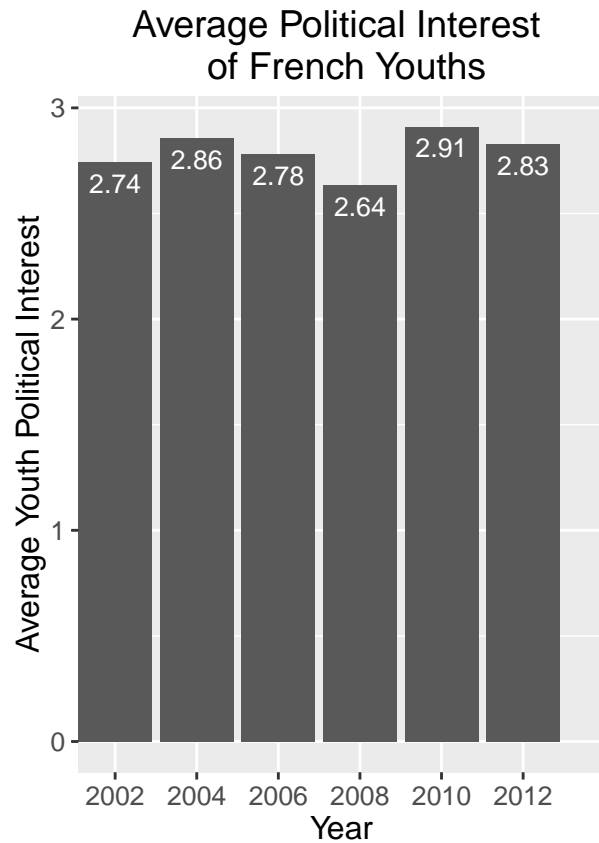
Further, after looking into the construction of different variables and the structure of the responses in the ESS, we keep only those variables that have a sufficient number of observations for our purposes and are directly relevant to us. Since many of the ESS variables are categorical and contains coded response levels, we drop categories like “No answer”, “Refusal”, “Don’t know”. We thus, collapse the categories in these variables in line with our intended research question.

As we want to merge the datasets by time as well as country, we had to undertake the process of equalising the both these variables. We created a new variable by the name ‘cntry’ in the Eurostat dataset which carried the country codes for each country included. This was required as the ESS data only contained the country codes while Eurostat data only showed the full names of countries. The time dimension was equalised as described above.

Lastly, ESS data was grouped by year and country after which the means of individual-level observations in the ESS data were taken within the designated groups. Hence, we now have average value of each variable for each country by each year covered. This new data frame (built from the original ESS data) was merged with Eurostat data to give us our final data frame containing the measure of youth unemployment and various indicators of political participation per country per year.

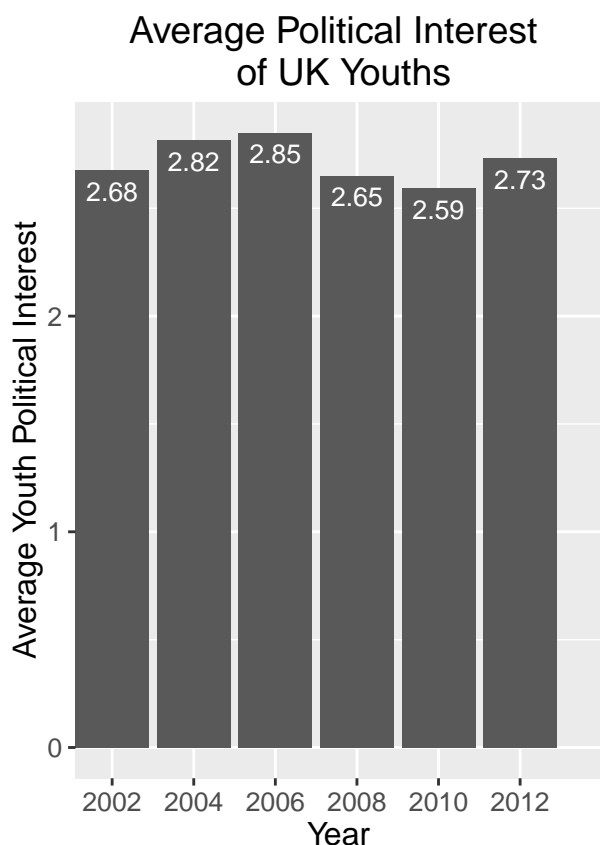
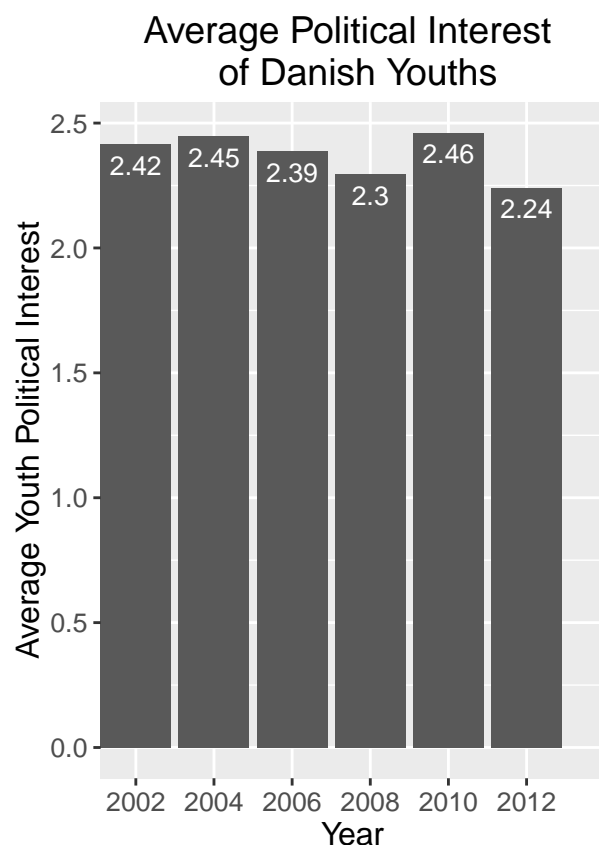
Descriptive Statistics

To start out with, it would be interesting to get an idea how much a given ESS category tends to vary from round to round for a single country. For example, here is a side-by-side comparison of how the average political interest of youth in France and Spain has varied between 2002 and 2012 (unfortunately, there is no ESS data for Spain for 2012). For reference, the scale here is 1 to 4, with a 1 representing someone who says they are “very interested in politics” and a 4 representing someone who says they are “not at all interested in politics.”



Though there are not drastic changes between survey years, for France we can still notice a small increase in political interest in 2008, and a marked decrease between 2008 and 2010, which is particularly interesting timing given the economic downturn that happened in 2008. With Spain, we can see that overall, there seems to be slightly higher political apathy on average (with the values veering closer toward 4, meaning no political interest whatsoever), and the values do not seem to have reacted noticeably to the economic downturn.

For further insight, we can look at the data for two other countries, Denmark and the UK.

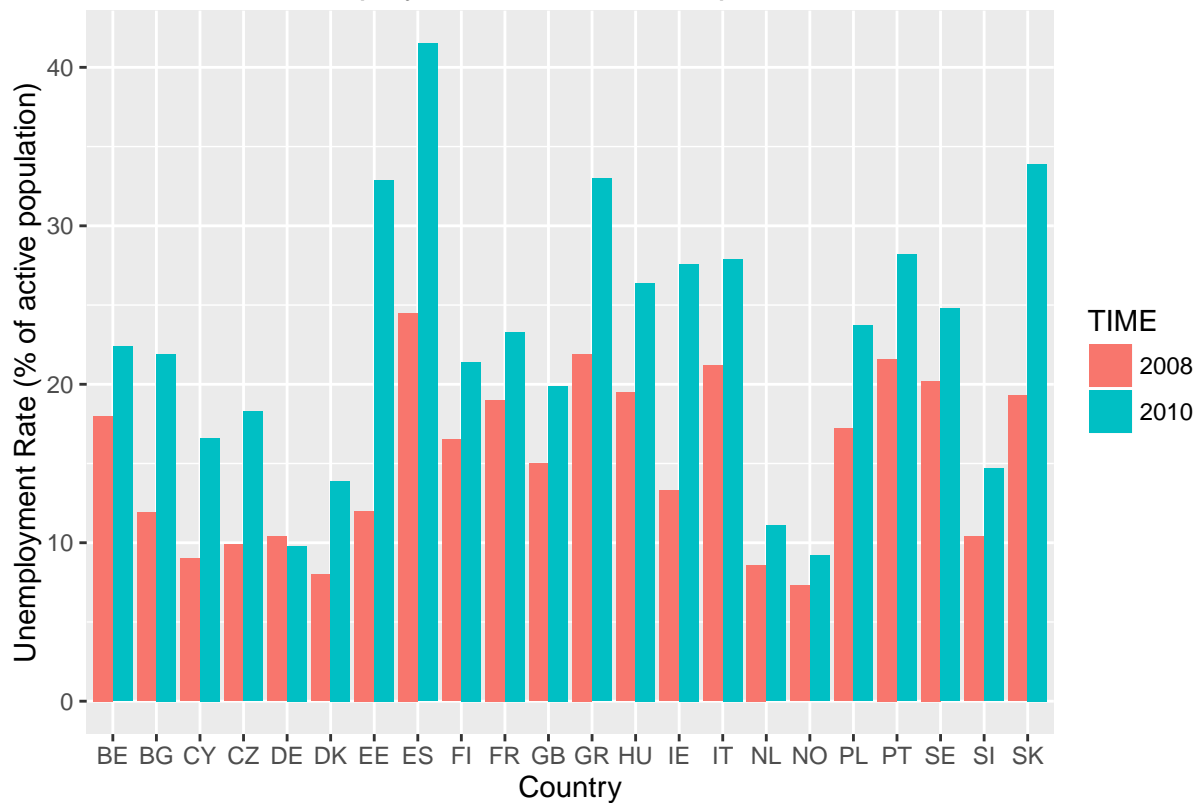


We can see that Danish youths are somewhat more politically interested than Spanish or French youths, but they still experienced a decrease in political interest from 2008 to 2010. However, such an effect is less noticeable for the UK, and the numbers even seem to suggest a very slight increase in political interest from 2008 to 2010.

It is of course important to note that political interest is just one metric for evaluating political engagement, and we intend to look at more indicators for our full research report. This is just meant to give an overview of what kind of variation in these metrics we can expect.

On the topic of youth unemployment, most countries experienced a significant increase in the unemployment rate from 2008 to 2010. The following table looks at this change for all European countries that exist in both datasets.

Youth Unemployment Rates in Europe, 2008 – 2010



All countries experienced at least a modest increase in the youth unemployment rate between 2008 and 2010, with the exception of Germany.

To start to bring the two concepts of youth political engagement and youth unemployment together, we can already run a simple bivariate regression using the data we currently have available.

<i>Dependent variable:</i>	
avgpolintr	
Value	0.012*** (0.003)
Constant	2.599*** (0.072)
Observations	81
R ²	0.134
Adjusted R ²	0.123
Residual Std. Error	0.248 (df = 79)
F Statistic	12.243*** (df = 1; 79)

Note: *p<0.1; **p<0.05; ***p<0.01

It indicates that increases in the youth unemployment rate in a country lead to a decrease in interest in politics (recall that a higher number here means a lower interest). Though the effect is significant, it is quite small. For the final project, we will investigate what kinds of control variables we could make use of to strengthen the regression.

Moving Forward

Analyzing the relationship between youth unemployment and the political engagement of young people would require us to think about the possible control variables that we need to include in our model. We will potentially be using at a third dataset from where we can extract variables that capture political stability, education levels, levels of employment related public protection, etc.

Since our current dataset only contains data up to ESS Round 6 (2012), we also intend to addend information from the 2014 ESS Round 7 to our existing ESS dataset. Another issue that we would need to deal with is how to use the different measures of political engagement from our current dataset in our analysis, since in this paper we only looked at one possible measure.