

Lab 6: Voting in Georgia

David Sichinava

29 2019 .

Introduction:

Follow the assignment step-by-step. Name your *.rmd* file using the following naming convention: *your_surname_lab6.rmd*.

Background:

In this exercise, we are going to analyze voting patterns in Georgia. Specifically, we will explore how people in Georgia were voting for the United National Movement party from 2008 to 2016. We will analyze how voting is correlated with socio-demographic characteristics. Download raw dataset [from here](https://goo.gl/QoCMrD) or type the following address in your browser: <https://goo.gl/QoCMrD>

Describe your data

Create a new notebook and import data. Name your data table as *geo*.

You can find a description of your variables in table 1. Each record in the table corresponds to a specific electoral district. There are information about 2008 presidential and parliamentary elections, 2012 parliamentary elections, and 2013 presidential elections data.

. 1: Variable description	
Variable	Description
UNMProp	UNM support in % in district
TurnoutProp	Turnout
Urban	% of urban population
Georgian	% of ethnic Georgians
Orthodox	% of orthodox christians
HiEd	% of people with higher education degree
WHC	% of white-collar workers
District	District ID
Elections	Elections

Data description

First, let's explore our data. Calculate mean proportion of UNM votes and mean turnout per each election. Which election has on average the highest value?

Now, analyze other variables. Is there any association between the proportion of white-collar workers and people with higher education degree? Why is this a case?

Use *ggplot2* library and construct a graph visualizing relation between education and employment type. Below you can find an example of this.

```
ggplot(geo, aes(x=HiEd, y=WHC))+  
  geom_point(aes(colour=Elections))+  
  geom_smooth(method="lm")+  
  labs(title="Higher Education and Employment",  
        x="Proportion of people with higher education",  
        y="Proportion of people with white collar jobs")
```

Calculate the same graph for visualizing relationship between UNM votes and electoral turnout.

Data analysis

Calculate correlation between UNM votes and turnout. How would you explain this pattern?

How to submit my assignment?

Upload zipped file to [this link](http://bit.ly/2Oy1QZB) before the start of our next meeting. You can also access the same link via browser by typing: <http://bit.ly/2Oy1QZB>