

Project I -2020-2021 Part time

The data refer to the elections for the candidates of both parties (Democrats and Republicans) for the presidential election of 2016. The data are given in two files

The one file has the socioeconomic characteristics of the counties and the second one the votes for the candidates.

We want to create a model using as response whether Clinton won over Sanders for the Democrats, using as explanatory variables the socio-economic characteristics of the counties.

We are interested to explain the behavior not to predict. You need to find a reasonable model selecting covariates and being able to use the model to explain the behavior of voters.

Description of the variables is given in separate sheet in excel.

Provide a report in pdf format with your findings. Be as detailed as possible so as your report to be self-explained. Explain why you selected the specific model, how good you think it is and any limitations that may apply.

You have also to upload in a separate file the R code you have used for the project.