

Swiss Political Survey Dataset

Christian Tresch
Mateusz Paluchowski
Nicolas Rabany
Matthias Tsa

Dataset description

Dataset is consisting of ~500k rows where each row represents one answer for political survey consisting of 75 various questions where answers belong to one category [no opinion; strongly disagree; disagree; agree; strongly agree]. Dataset is sparse where not many people answered all questions thus it will require preprocessing and appropriate NaN handling.

Graph construction

Each person that answered to the survey will be defined as a node of the graph. Their connection weights will be modelled as the thresholded L2 distance between their responses to a selected subset of questions. These subsets will be predefined by hand according to the more general topic to which the question belongs to. This will result in different graphs depending on the question topic chosen to define the connection weights.

Problem description & questions to answer

Which political party are the people from/for? We want to guess the political party of the people by analysing the answers to the survey. Find if the number of political parties reflect the amount of possible clusters within the dataset - if all people are correctly represented in the parliament by some party.

