Voting Advice Application - Danish data

Ann-Sofie Hansen, Dennis Hansen, Oskar Harmsen & Susanne Sundgaard

The data

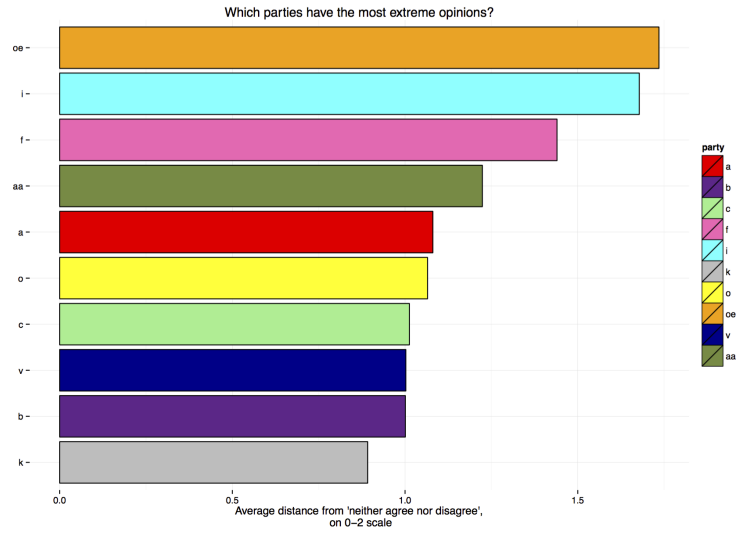
We have scraped the responses from the Danish politicians to the voting advice application (VAA) at DR's homepage. 724 candidates answered the VAA out of a total of 799 candidates running (source) and thus accounts for 1.630.587 personal votes out of 1.762.656 given. This difference can to a large extent be attributed to the fact that two opposing party leaders, Lars Løkke Rasmussen and Helle Thorning-Schmidt, did not partake in the VAA. The 1.630.587 personal votes accounts for 46,3% of all (valid) votes (3.518.987 in total).

The candidates are asked to rank 15 questions on political issues on a "Totally disagree-Mostly disagree-Indifferent-Mostly Agree-Totally agree" scale. The questions vary from tariffs on cigarettes over public sector growth to the amount of influence given to EU. The questions are all weighted equally and all 724 candidates have answered all questions.

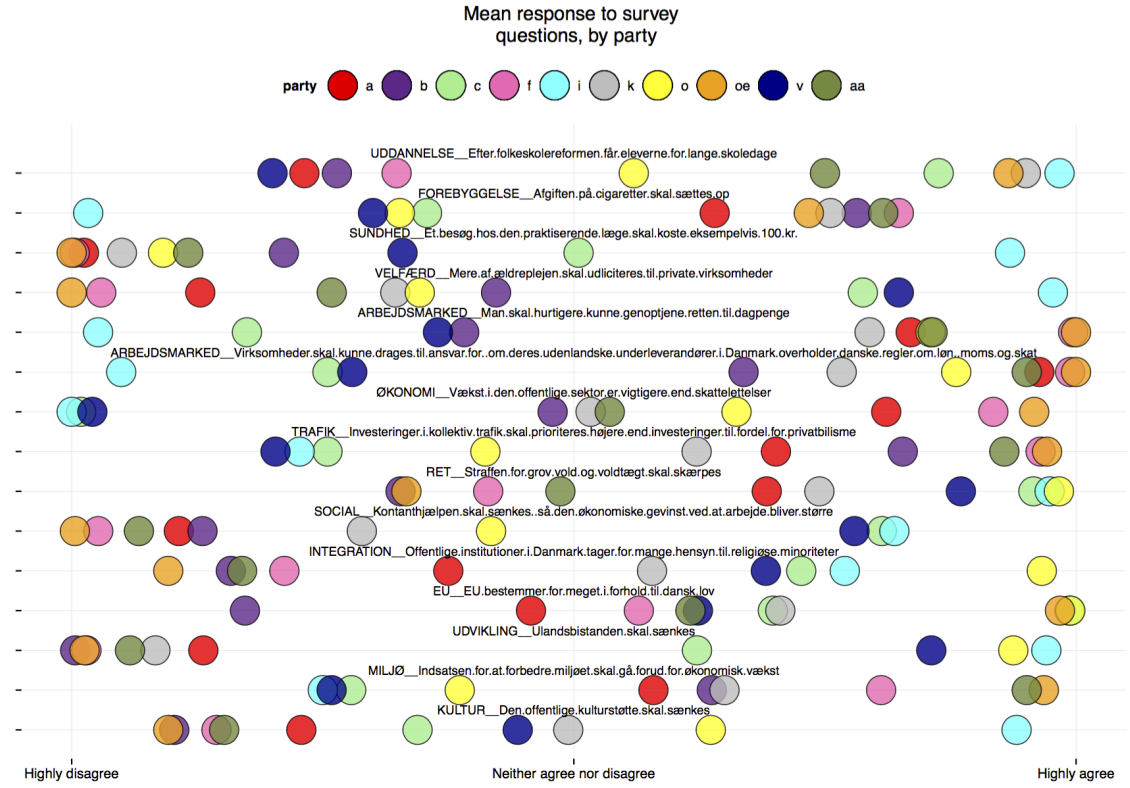
Furthermore we have scraped additional descriptive data on the candidates such as gender, age, current position, whether they ran for office at the previous election and whether they were elected at this election or not. There are five candidates who are not running for a particular party (independents) and these are discarded since they do not recieve a remarkable amount of votes.

Descriptive statistics

We start by computing an overview of the mean responses to question on a party-level. (Appendix??) First of all we see that the mean responses to all questions seem rather spread out in the sense that there is no immediate clusterings around one end of any questions. Second of all we see a strong tendency for a pattern at least for the orange and turqoise dots (Enhedslisten and Liberal Alliance) that seem to linger around the (opposing) edges of the mean response spectrum and rarely around the middle. There are also some dots that are almost completely, or to a large extent, overlapping - as is often the case for Enhedslisten (orange) and SF (pink), which makes sense politically. However we also see opposing parties joining views in matters of EU (Enhedslisten and DF) and the public school reform (Enhedslisten and Liberal Alliance).



These observations are validatet even further when we compute the average distance from the "neither agree nor disagree" for each party (figure 2), where Enhedslisten and Liberal Alliance, not surprisingly cf. figure 1, are distinctively more extreme in their opinions. The average distance for the most extreme (Enhedslisten) is almost twice as large as the least extreme (Kristendemokraterne).



Principal Component Analysis

A Principal Component Analysis is a method within unsupervised statistical learning for reducing dimensions in a parameter space. The VAA-data is really a 15-dimensional parameter space but by using PCA we can retain the two dimensions that explain the most variance in the data. [mere teknisk om PCA] When plotting the two first components (i.e. the two most influential) we see the distribution of the parties