Swiss Political Survey

Mateusz Paluchowski, Nicolas Rabany, Christian Tresch, Matthias Tsai

Problem description

- Which political party are the people from?
- What are the relationships between parties / cantons?
- Is there a party missing?

Data collection

Political opinion survey from SmartVote

Full Dataset :

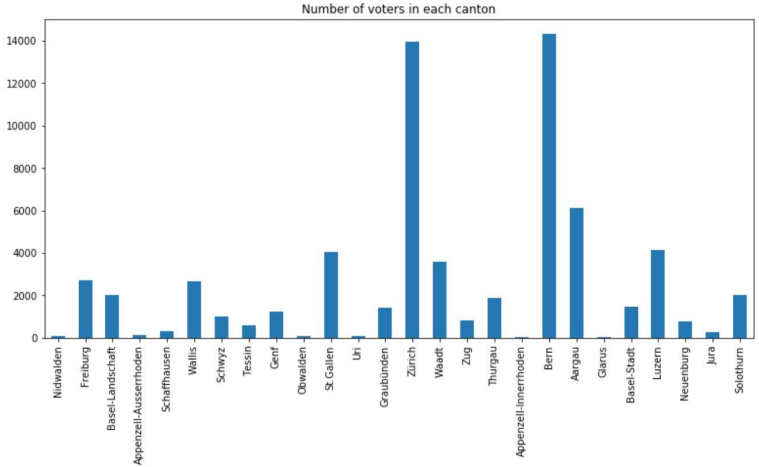
- 500’000 voters and 4’000 candidates
- Content:
  - 75 questions + personal informations
  - Party of the voters recommended by Smartvote
  - Party of the candidates
- Possibility to weight each question

Cleaned Dataset:

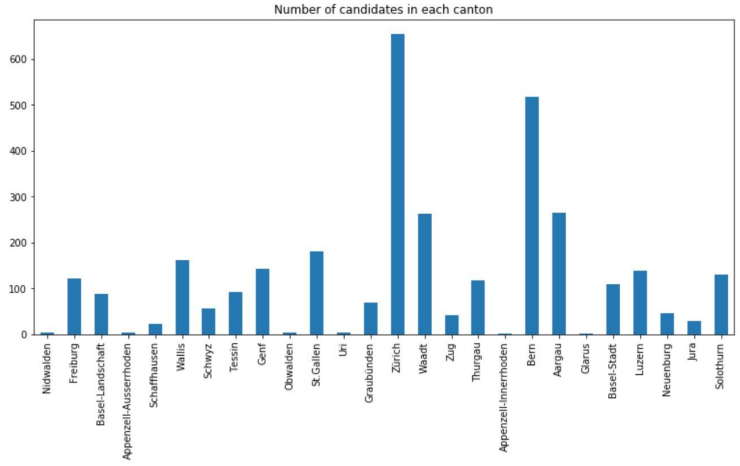
- People who answered all the questions: 66’000 voters and 3’200 candidates
- Content:
  - Questions
  - Cantons
  - Parties

Exploration

Geographical distribution

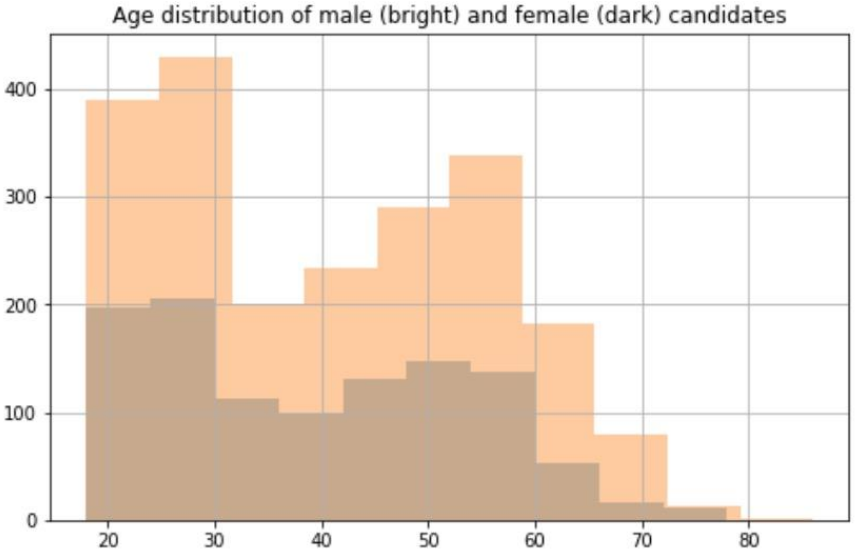


Voters

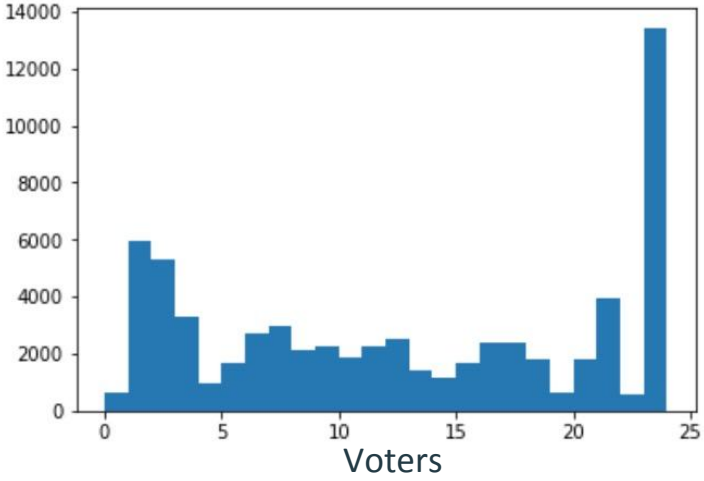


Candidates

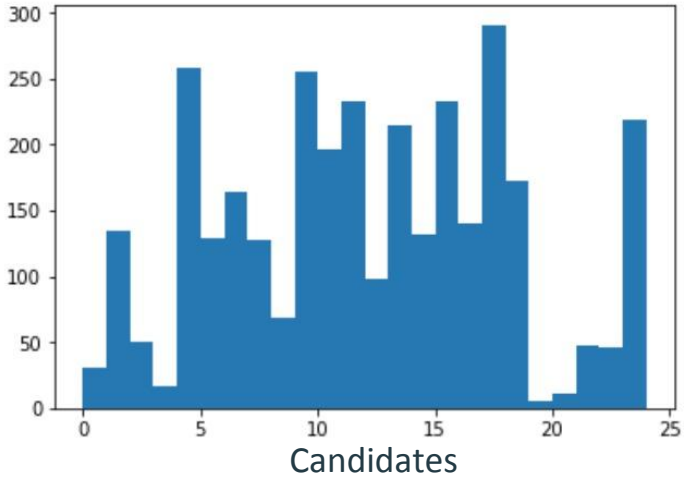
Age Distribution



Political Party Distribution



Voters



Candidates

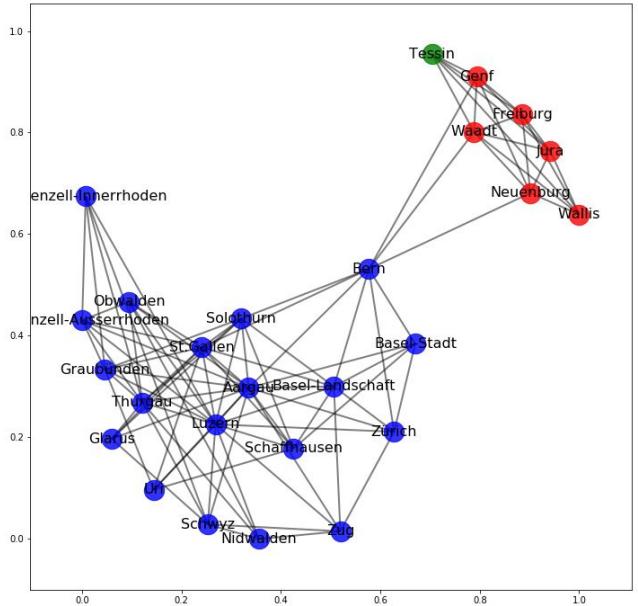
Exploitation

Graph construction

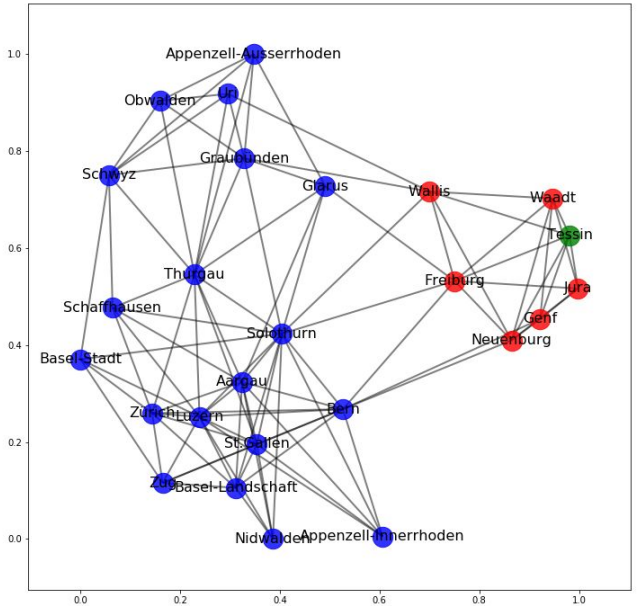
1. Compute the **party recommendation** for the **voters** (using cosine distance)
2. Compute **party and canton centroids** over the voters and candidates datasets
3. Use cosine distance to compute the **similarities between centroids**
4. Construct the **weight matrix**
5. **Sparsify** the matrix: keep only the **5 strongest edges** between the nodes

Canton graphs

- French part
- German part
- Italian part



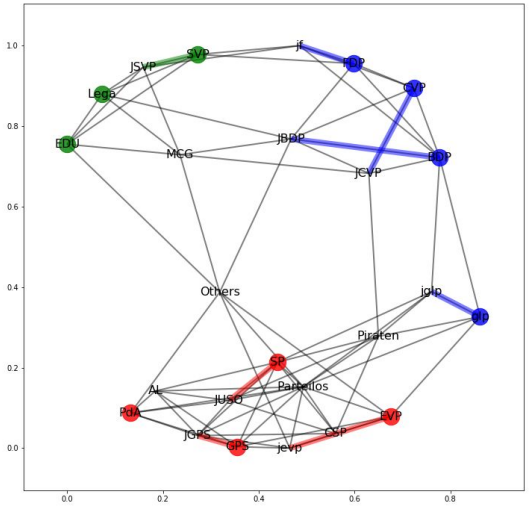
Voters



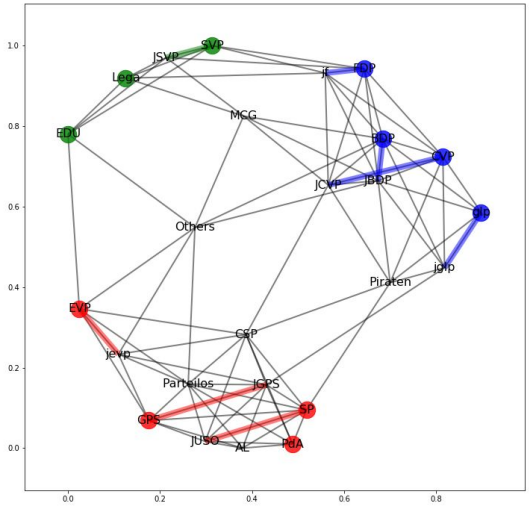
Candidates

Party graphs: individually

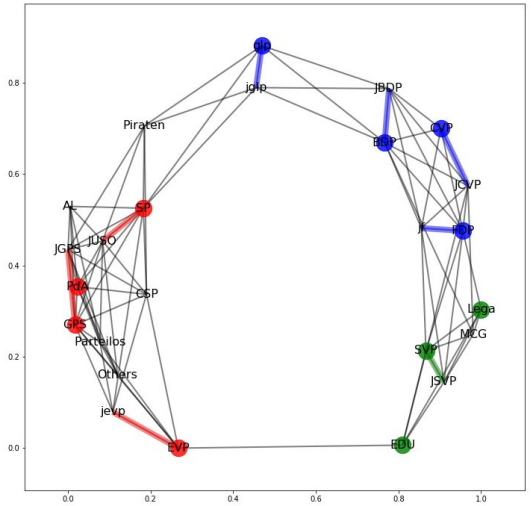
- Left-Liberal
- Right-Liberal
- Right-Conservative



Smartvote recommendations for voters

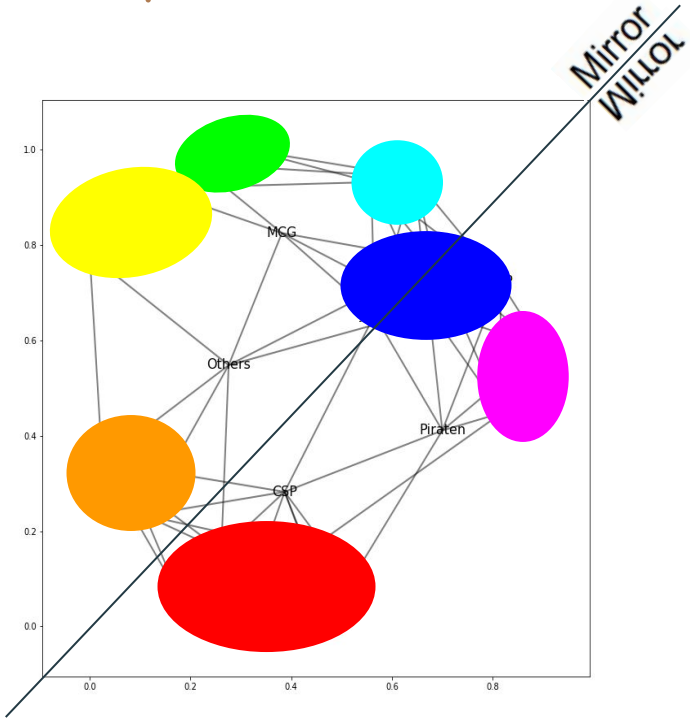


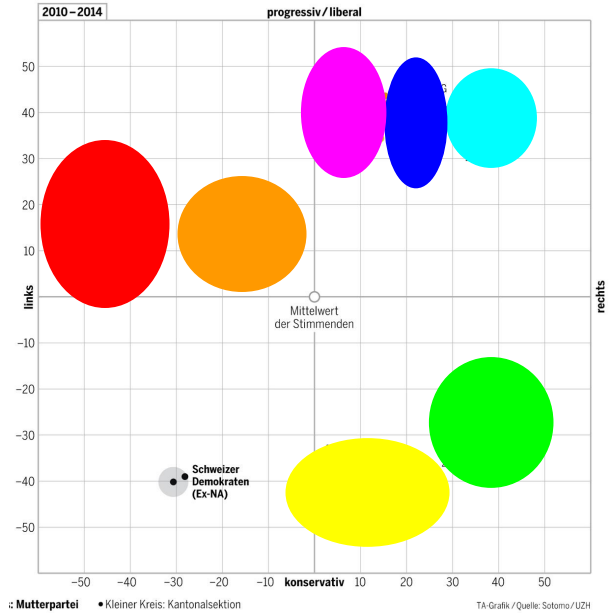
Our recommendations for voters



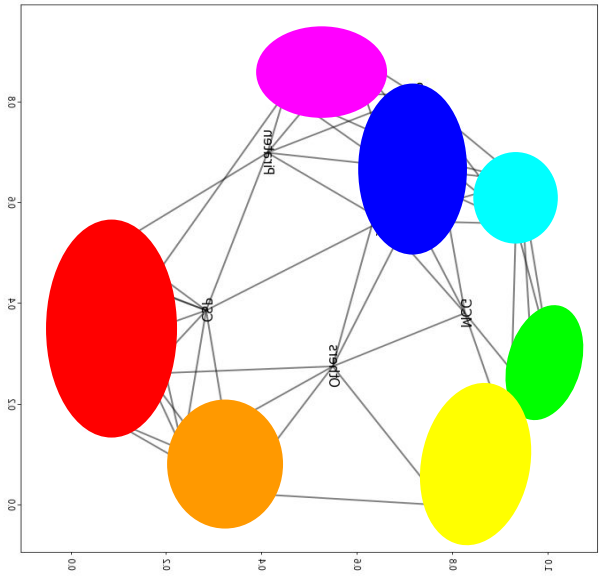
Candidates

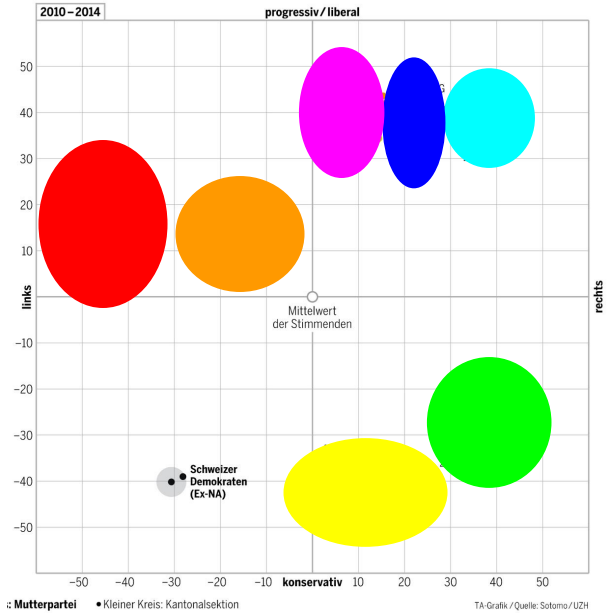
Liberal/Conservative - Left/Right Coordinates





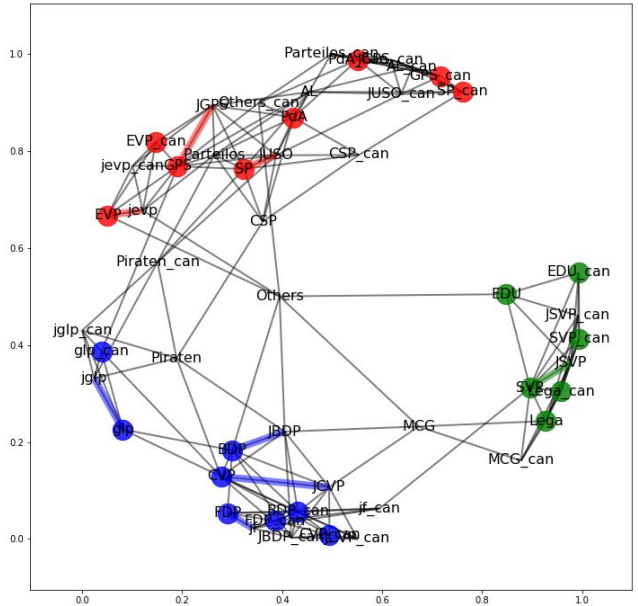
Liberal/Conservative - Left/Right Coordinates



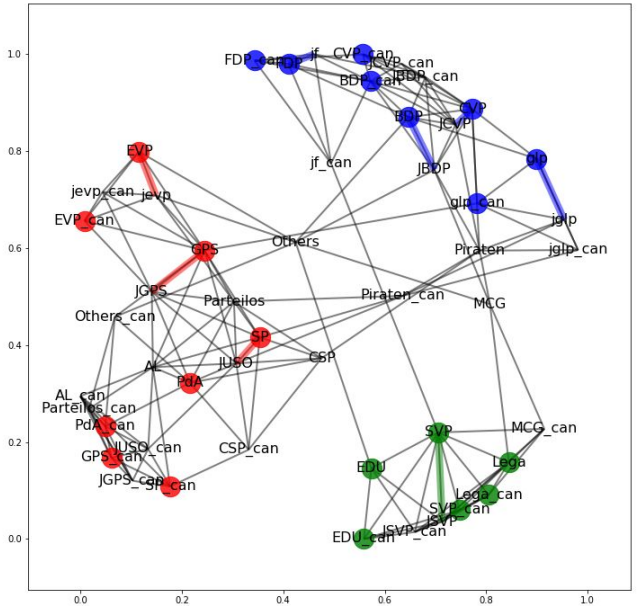


Party graphs: combined

- Left-Liberal
- Right-Liberal
- Right-Conservative

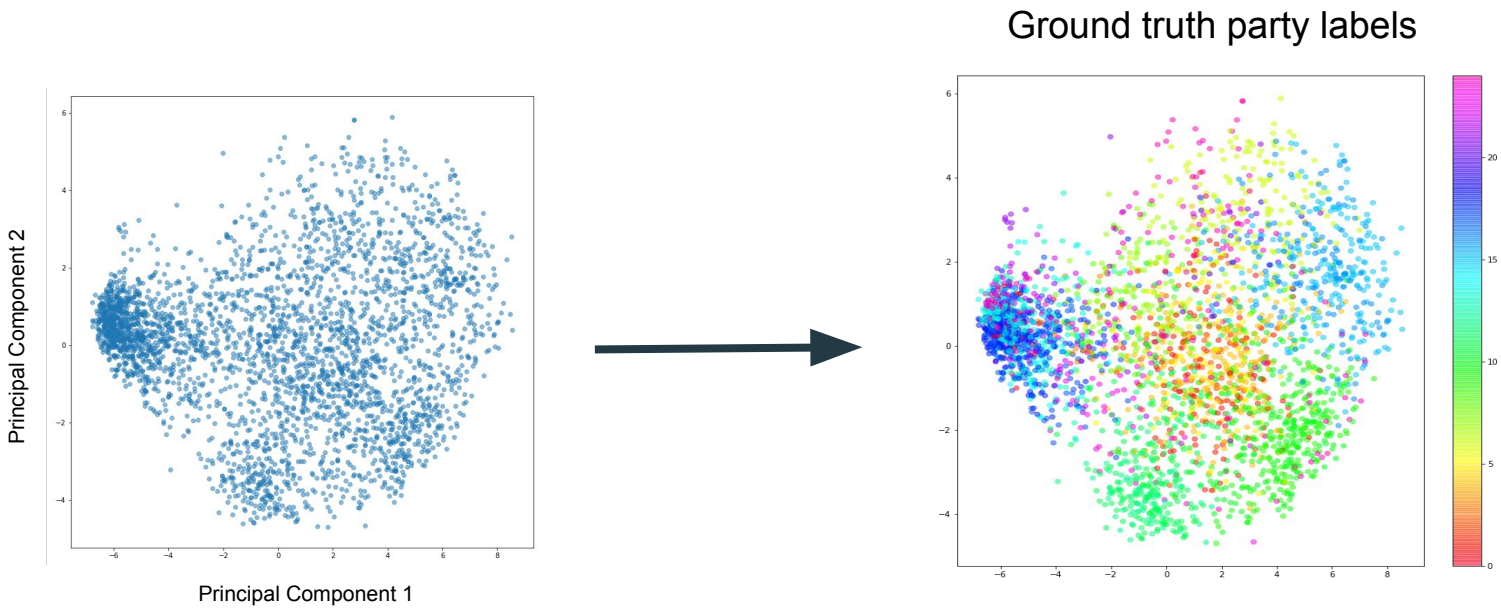


Candidates and voters (SmartVote recommendations)

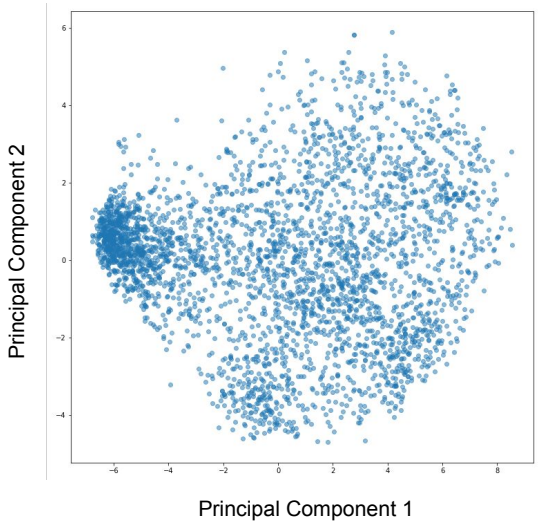


Candidates and voters (our recommendations)

PCA on Candidate Dataset

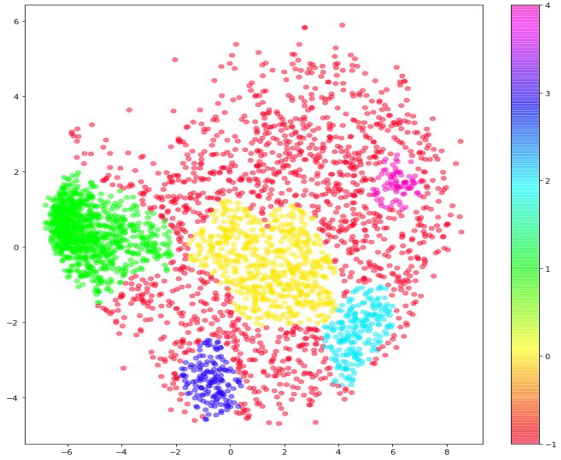


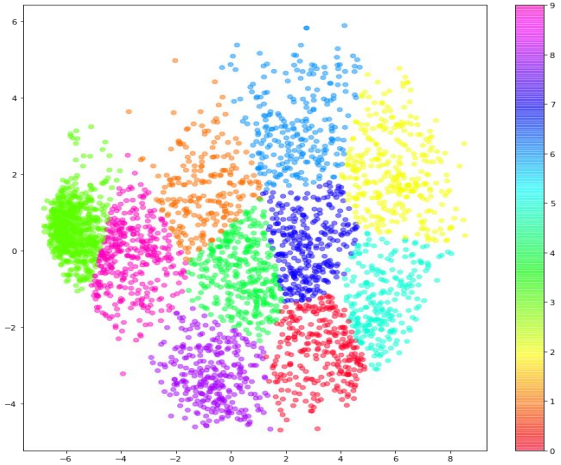
PCA on Candidate Dataset



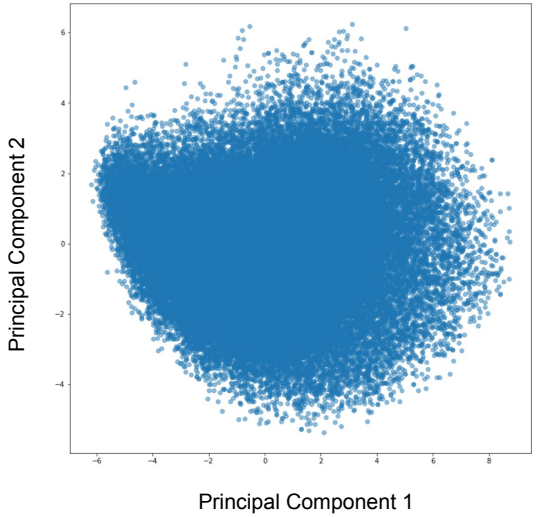
DBscan

K-means



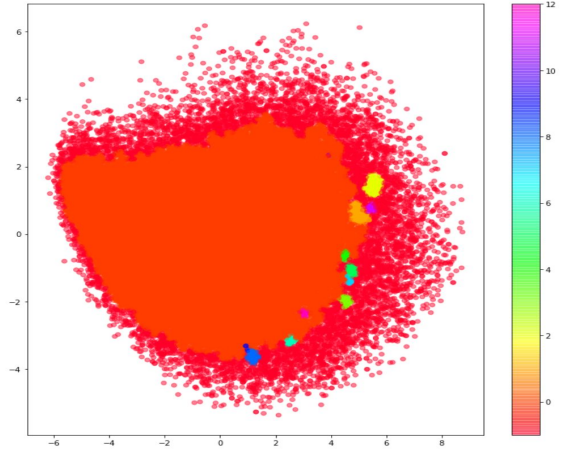


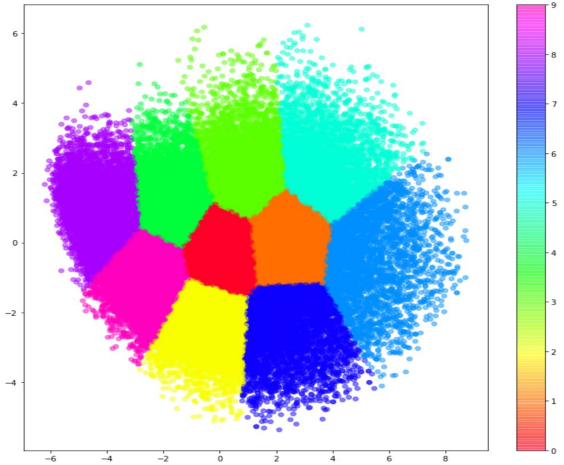
PCA on Voter Dataset



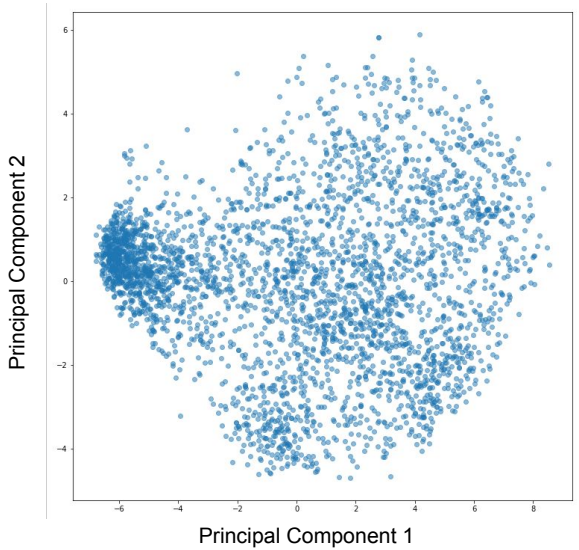
DBscan

K-means





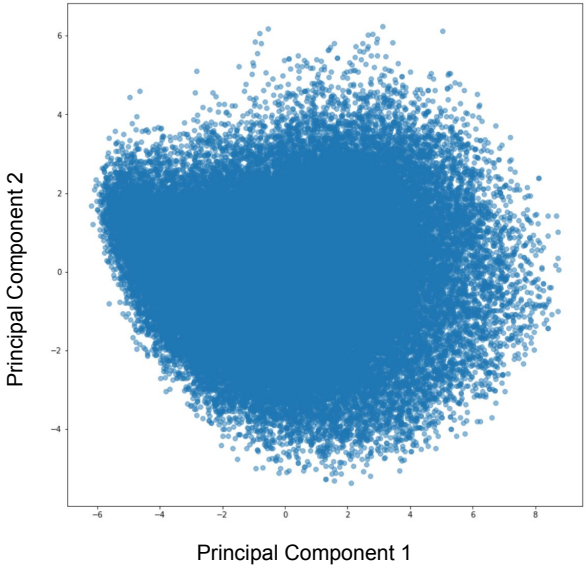
Gaussian Mixture Model on Candidate Set



GMM

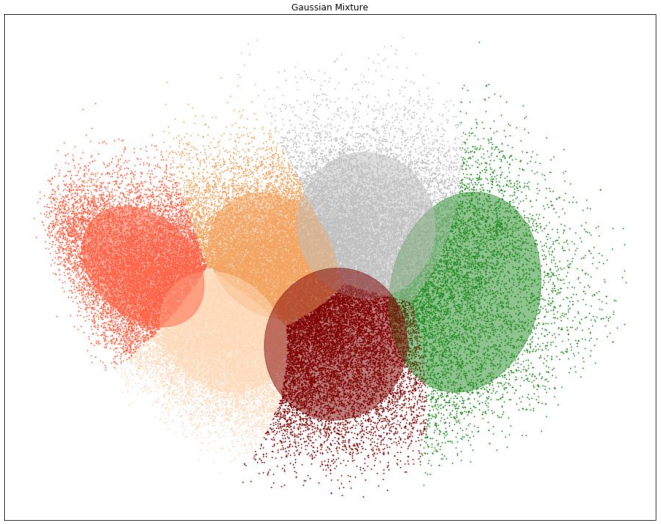
GMM with 6 components

Gaussian Mixture Model on Voter Set

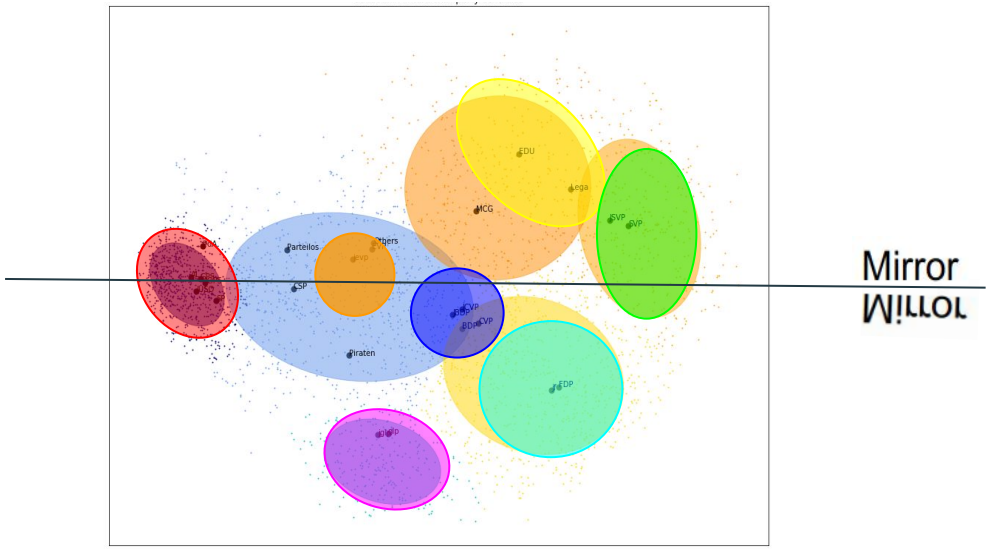


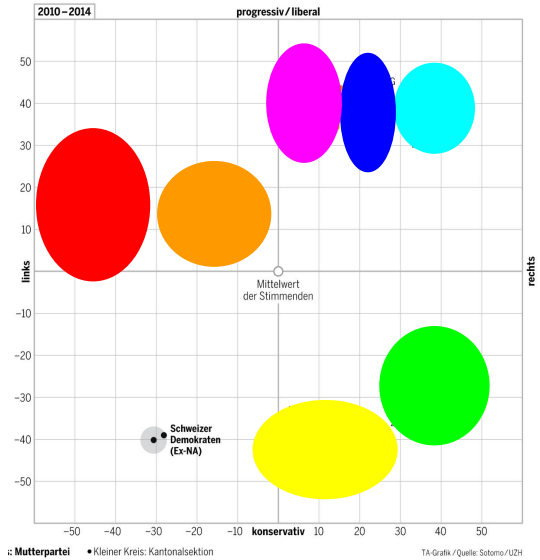
GMM

GMM with 6 components

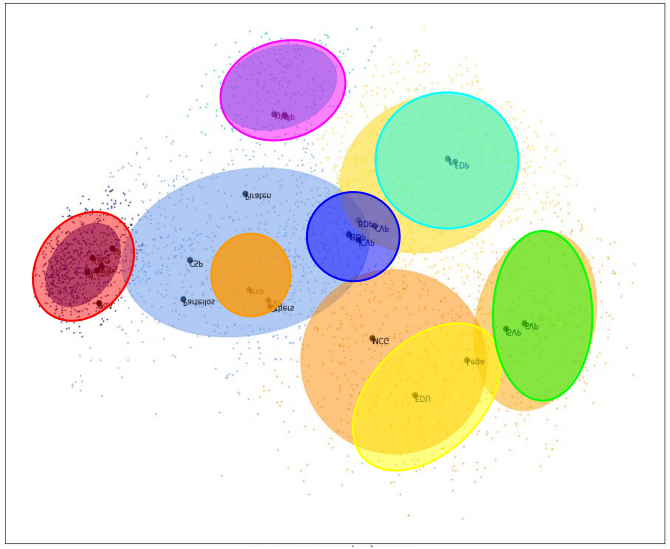


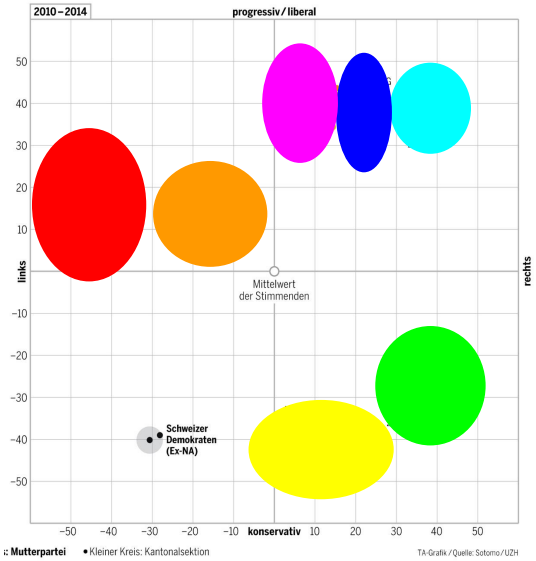
Gaussian Mixture Model Evaluation





Gaussian Mixture Model Evaluation





Conclusion

Graph-Based Analysis

- Proved that very similar party recommendations could be achieved from our cosine distance based metric as the one from smartvote.
- Managed to visualize the Röstigraben.
- Successfully grouped left/right and liberal/conservative parties without any supervision by building political party graphs.

PCA-Based Analysis

- Candidates are easier to separate than the voters.
  Candidates cluster easily, but voters have more homogenous distribution.

  -> Voters opinion is more uniformly distributed

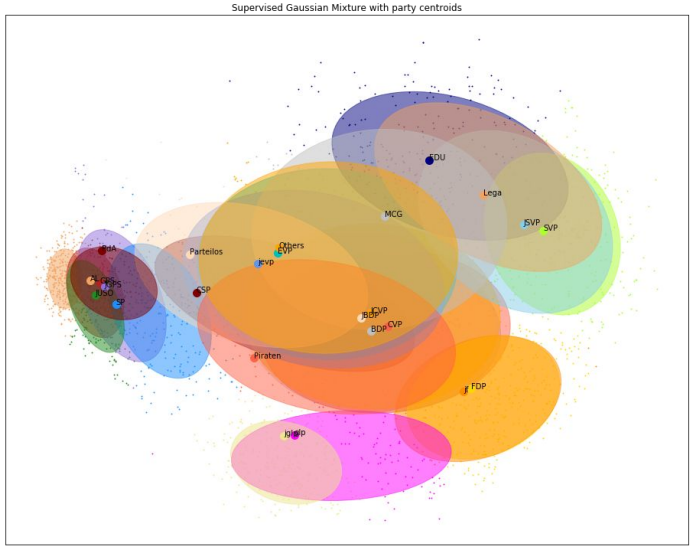
- We can not guess if there is a missing party. A missing party could be anywhere between existing parties.

- Hypothesis: Candidates can be subject to a bias in their opinions

Thank you!

Supervised Gaussian Mixture Model

Training on candidate data set
(GMM with 25 components)



Supervised
GMM

Prediction on voter dataset
(GMM with 25 components)