

# Master thesis proposal

## Analyzing MAD analysis of Poland's 2025 presidential elections

Krzysztof Bartoszek

September 27, 2025

### **Background — evolution models**

Following Poland's 2025 presidential elections there were claims of irregularities in a large number of voting stations [1]. This analysis, based on the outlier detection by the mean absolute deviation (MAD) approach was heavily criticized, e.g., not performing a symmetric analysis for both candidates in the run-off or doubts about the grouping of voting stations. However what was missing was a analysis of this analysis that takes into account voter flow or a comparison to an empirical background distribution based on past election results. The latter is crucial for checking if grouped voting stations do actually form a homogeneous group with respect to the results.

### **Thesis project**

The first step of the thesis will be to implement an algorithm to recreate the grouping of the voting stations as described in [1], as these are not publicly available. Then, after identifying flagged voting stations with respect to *both* candidates investigate these stations with respect to voting station type and size, first round results, and voter flow. A next step would be to rerun the analysis on previous presidential elections, investigate whether the distribution of flagged stations was different in 2025, and check how voting station type and size, first round results, and voter flow affect flagging of stations

in the past. The comparison with the past would allow for testing if the grouped of voting stations are actually homogeneous.

## Goals

The below are possible general goals. Depending on the student they will be adjusted in the direction of the student's interests.

1. Does voter flow between the first round and run-off explain the flagging of stations?
2. Are the grouped stations homogeneous, what is the probability of obtaining a homogeneous grouping by the method?
3. Does stations size or type affect the chances of being flagged?
4. Is the distribution (with respect to e.g., size, type, first round results) of flagged stations the same for both candidates?

## Data

The project will be based on official election results from the Polish National Electoral Commission ([www.pkw.gov.pl](http://www.pkw.gov.pl)).

## References

- [1] K. Kontek. Reassessing poland's 2025 presidential runoff: Using spatially-grouped mad detection to recalculate the result. 2025. URL <https://ssrn.com/abstract=5296435>.