

Analiza poparcia politycznego i stosunku do kampanii prezydenckiej 2025 w oparciu o Reddita

Szymon Lato

June 2025

Wstęp

Celem projektu było zbadanie sentymentu i ocena emocji w związku z kampanią prezydencką 2025. Ocena została przeprowadzona na wpisach z portalu Reddit z subredditów:

- r/Polska
- r/PolskaPolityka

Źródło danych

Posty i komentarze pobrano za pomocą biblioteki **praw** i przefiltrowano pod kątem obecności słów kluczowych związanych z polityką i wyborami (nazwiska kandydatów, "wybory", "kampania" i tego typu zwroty).

Każdy komentarz został przypisany do jednej z trzech faz kampanii wyborczej w zależności od daty jego publikacji:

- **pre** – przed pierwszą turą (do 18 maja 2025),
- **between** – między pierwszą a drugą turą (19–31 maja 2025),
- **post** – po drugiej turze (od 1 czerwca 2025).

Analiza sentymentu i emocji

W celu oceny nastrojów społecznych w komentarzach Reddita wykorzystano pięć modeli językowych. Część z nich operowała na oryginalnych komentarzach w języku polskim, inne na wersjach przetłumaczonych na język angielski (wykorzystałem GoogleTranslate).

1. PaReS – transformer polityczny (PL)

Model `eevvgg/PaReS-sentiment-political-PL`, dostrojony do analizy sentymentu w tekstach politycznych w języku polskim. Zwracał etykiety `[positive, neutral, negative]`.

2. NLPtown – BERT z oceną 1–5 gwiazdek (PL+multi)

Model `nlpstown/bert-base-multilingual-uncased-sentiment` służy do klasyfikacji sentymentu w pięciopunktowej skali (1-5 gwiazdek)

3. VADER (EN)

Słownikowo-regułowy model anglojęzyczny, obliczający tzw. wartość `compound`, która została przekształcona do etykiet sentymentu (pozytywny, neutralny, negatywny).

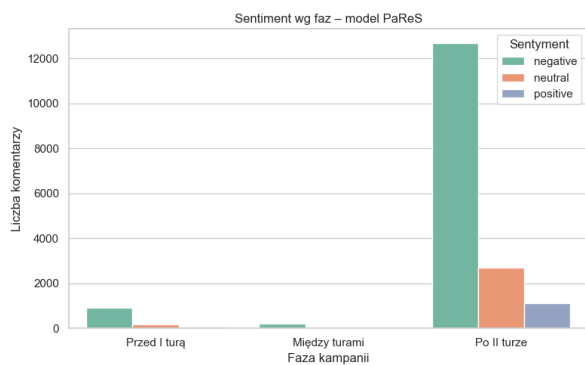
4. TextBlob (EN)

Prosty model o charakterze lingwistycznym. Oceniał `polarity` i `subjectivity`, a następnie przypisywał etykiety sentymentu.

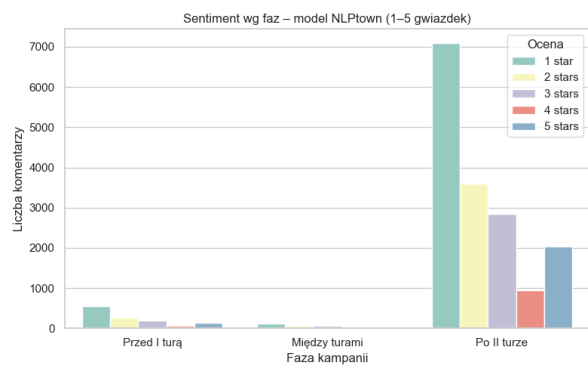
5. Text2Emotion (EN)

Biblioteka analizująca emocje wyrażone w tekście. Każdy komentarz został oceniony pod kątem pięciu emocji: *Happy, Angry, Surprise, Sad, Fear*.

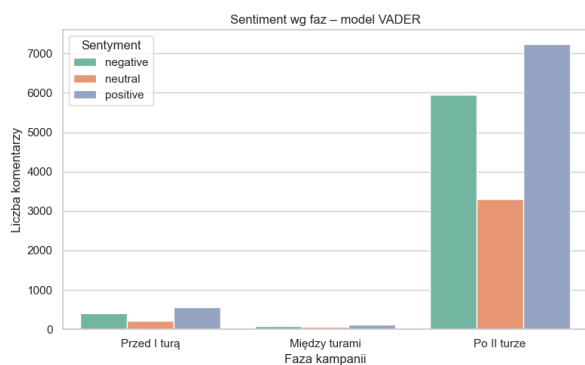
Porównanie wyników modeli



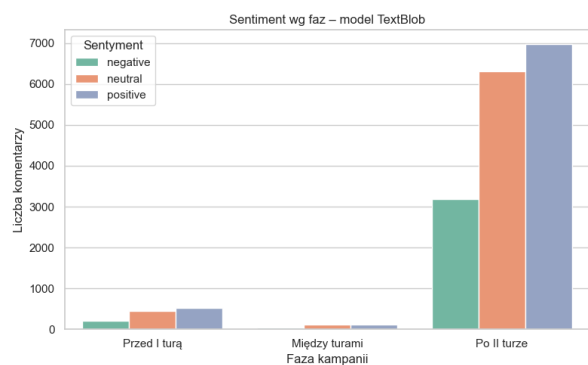
Model PaReS



Model NLPtown



Model VADER

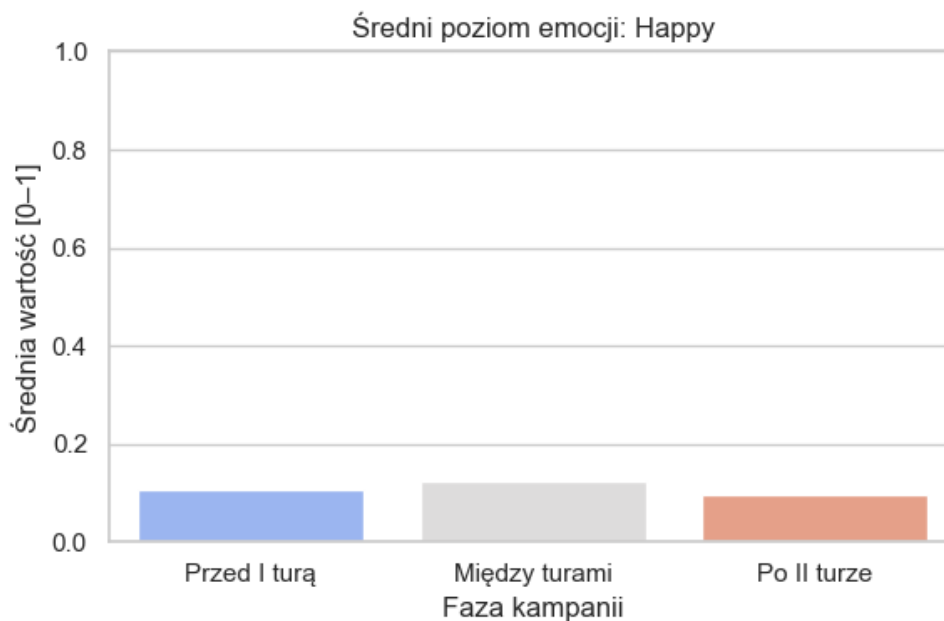


Model TextBlob

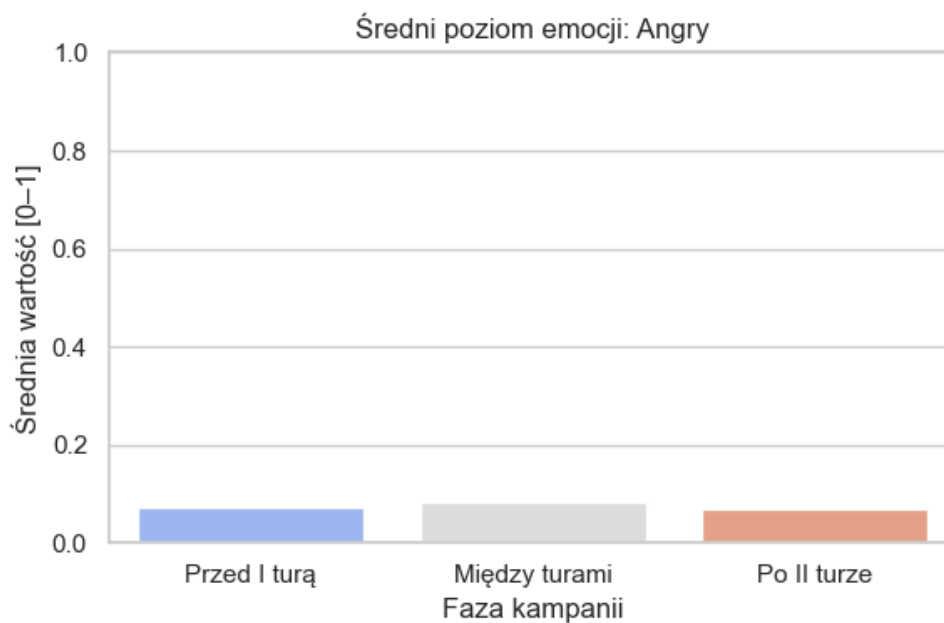
Rysunek 1: Porównanie rozkładu sentymentów w podziale na fazy kampanii wyborczej według czterech różnych modeli

Emocje według modelu Text2Emotion

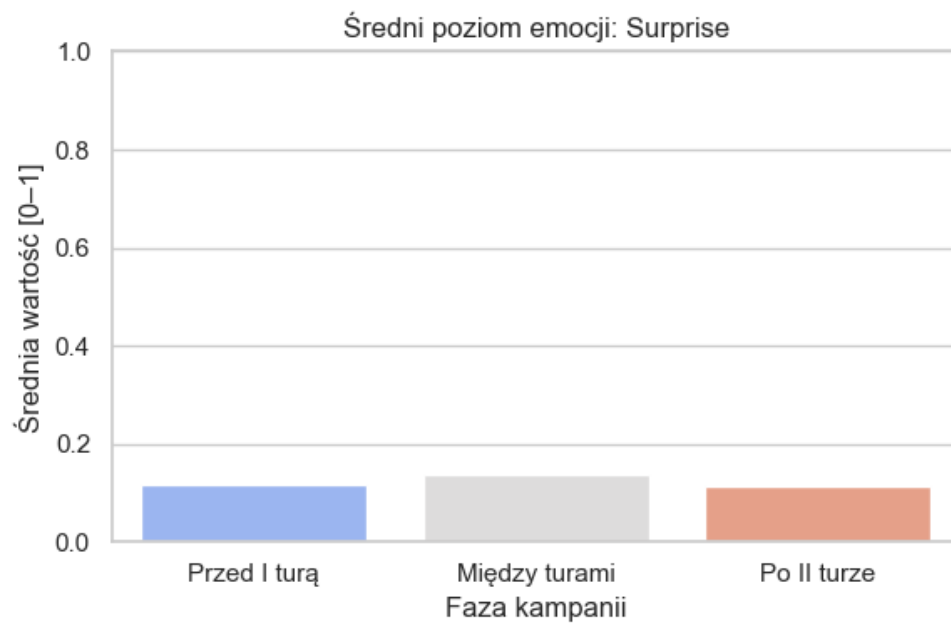
Poniżej przedstawiono średni poziom pięciu emocji wykrytych przez model Text2Emotion w zależności od fazy kampanii:



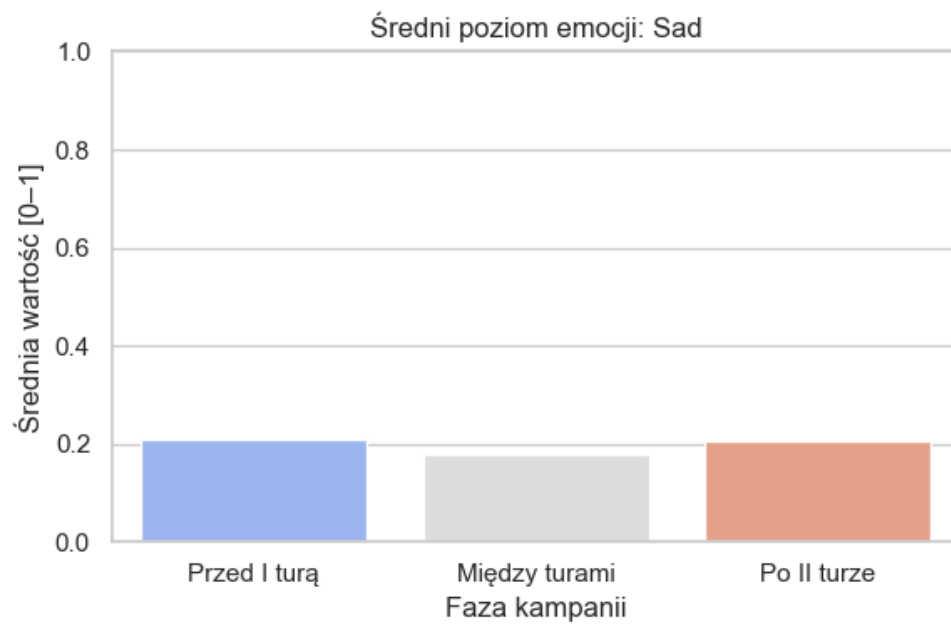
Rysunek 2: RADOŚĆ



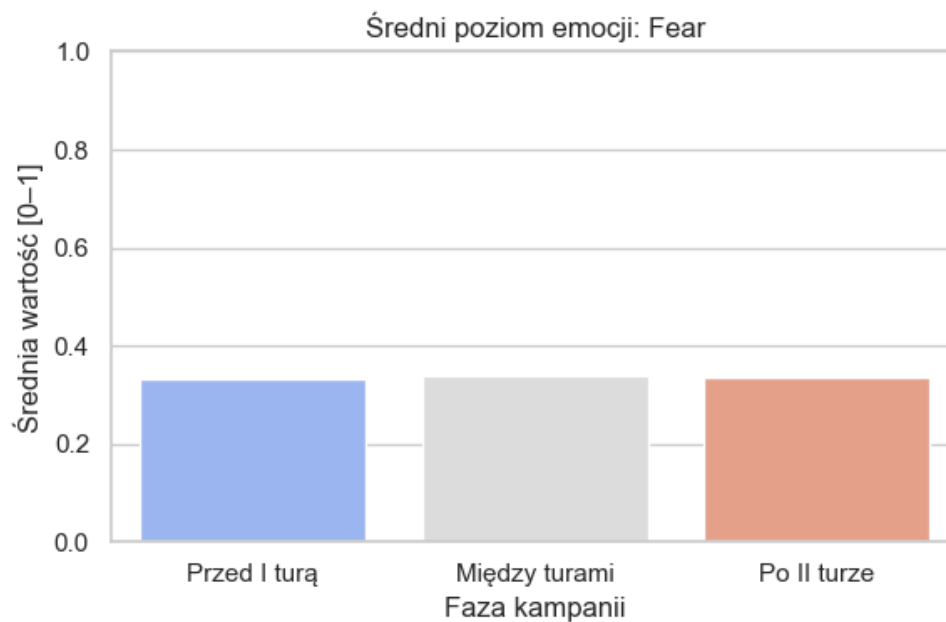
Rysunek 3: GNIEW



Rysunek 4: ZASKOCZENIE



Rysunek 5: SMUTEK



Rysunek 6: STRACH

Wordcloudy

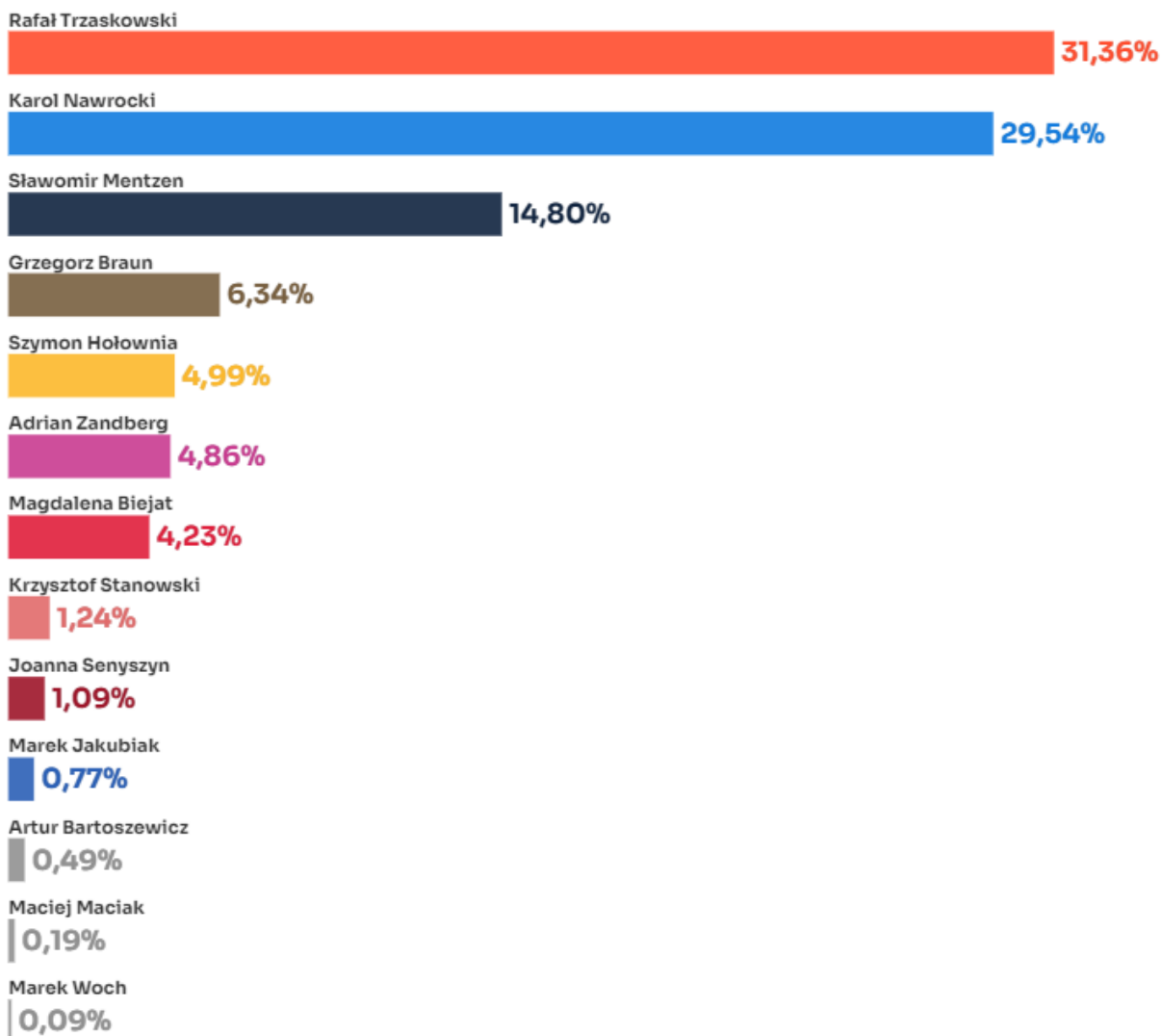
Najczęstsze słowa pojawiające się w związku z kampanią.



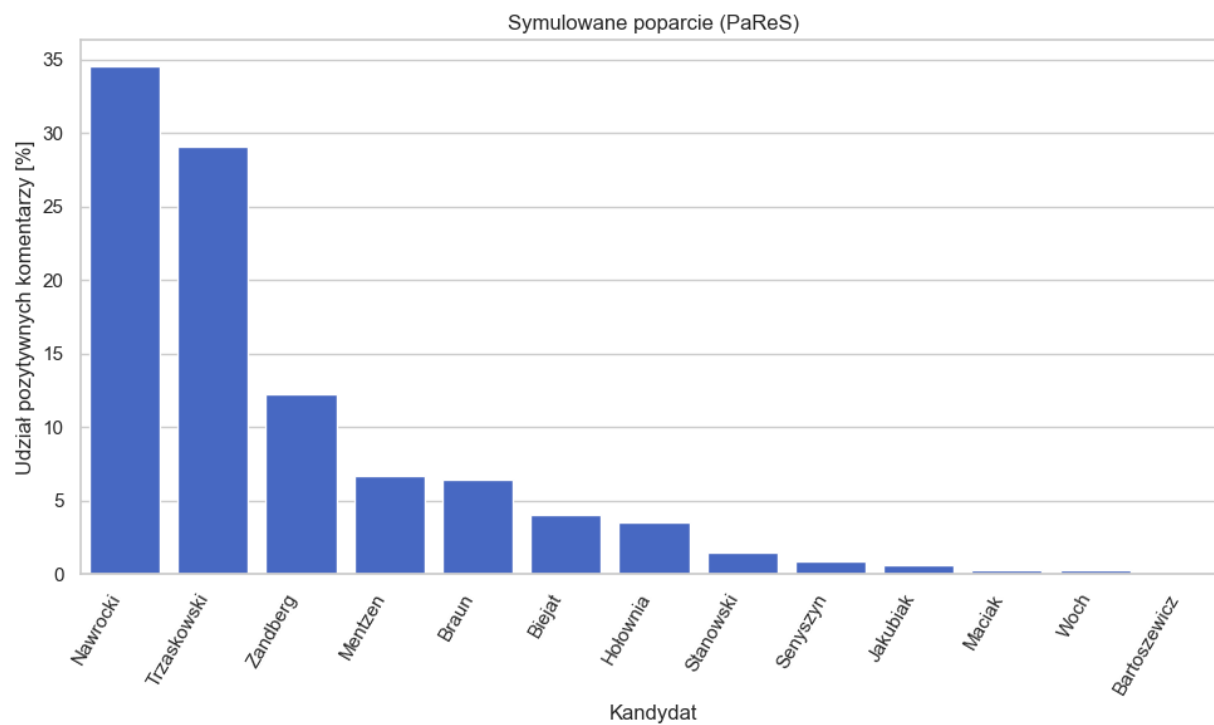
Rysunek 7: Chmura słów — faza przed pierwszą turą

Symulowane sondaże

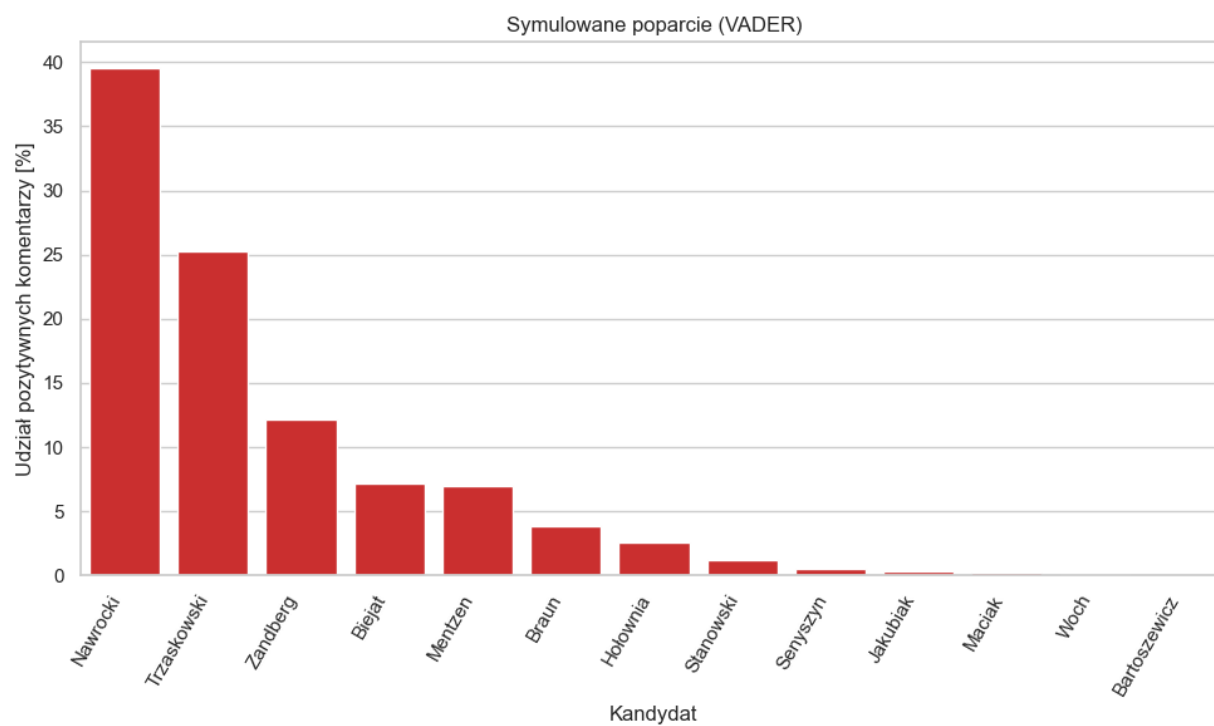
W oparciu o ilość pozytywnych opinii wygenerowałem symulowane słupki poparcia dla każdego z kandydatów. Każda opinia została według słów kluczowych przydzielona jednemu bądź kilku kandydatom. Potem zliczono liczbę opinii



Rysunek 10: Oficjalne wyniki i słupki poparcia dla pierwszej tury



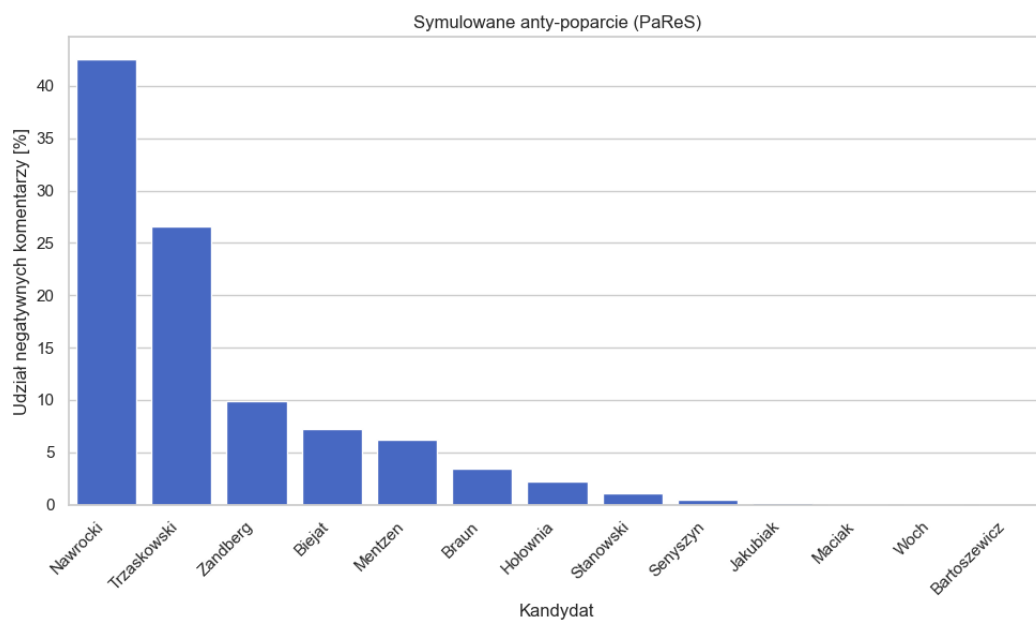
Rysunek 11: "Poparcie" wg PaRes



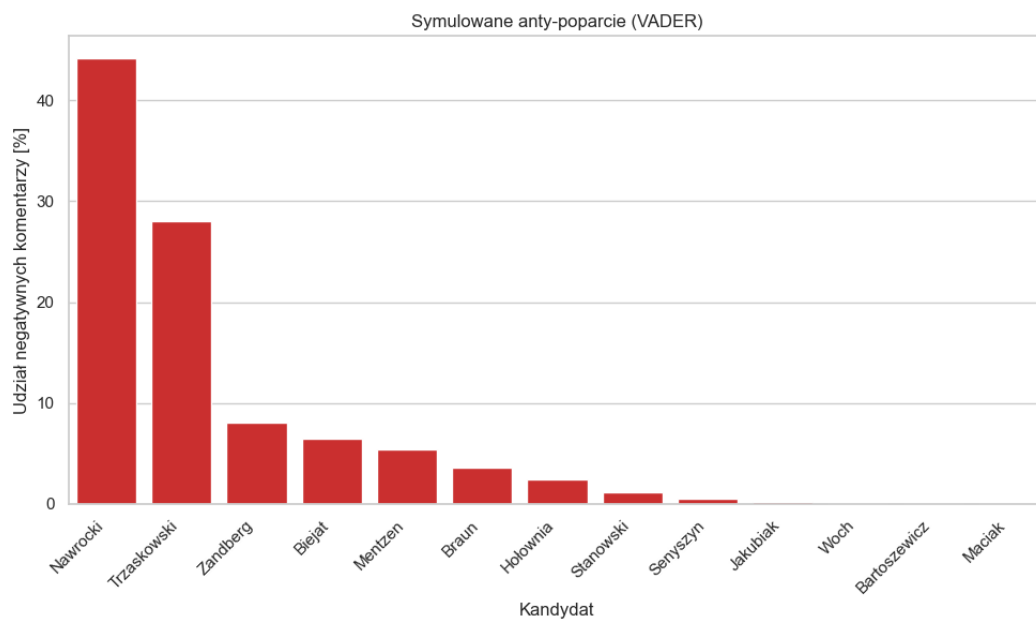
Rysunek 12: "Poparcie" wg Vadera

Anty-poparcie – słupki hańby

Podobnie jak w przypadku sondaży, tutaj na podstawie opinii negatywnych oceniamy, jak dużo anty-poparcia ma każdy z kandydatów.



Rysunek 13: Anty-poparcie wg PaReS



Rysunek 14: Anty-poparcie wg VADER