

# Ocena jakości sondaży przed wyborami prezydenckimi

Robert Konieczny

# Sondaże przed wyborami prezydenckimi

- Wybory prezydenckie:
  - I tura – 10 maja 2015
  - II tura – 24 maja 2015
- Sondaże i ośrodki demoskopijne:
  - CBOS
  - Estymator
  - GfK Polonia
  - IBRIS (Homo Homini)
  - Millward Brown (dawne SMG/KRC)
  - PP/ROBOP
  - TNS Polska (dawny TNS OBOP)

## Sondaże i ośrodki o niejasnej proveniencji:

- Pressmix
- Sonda uliczna ewybory.eu



# Po co sondaże?

- Dają nam wyborcom, obywatelom, społeczeństwu wgląd w nasze preferencje i pozwalają podejmować racjonalne decyzje (politykom też).
- Pozwalają lepiej rozumieć, analizować zachodzące procesy.
- Są jakąś kontrolą procesu wyborczego (vide rozjazd wyników exit poll i wyborów samorządowych i niniejsze sympozjum...)

# Konkurs o Puchar Pytii – Wybory Prezydenckie 2015

- I edycja
- Konkurs organizowany przez Centrum Badań Ilościowych nad Polityką UJ
- Kapituła:
  - Karol Życzkowski – Przewodniczący
  - Dariusz Stolicki – Sekretarz
  - Członkowie:
    - Jarosław Flis
    - Robert Konieczny
    - Zbigniew Puchała
    - Wojciech Słomczyński

# Pytia i wyrocznia delficka



Jacek Malczewski – Pytia

W świątyni Apollina w Delfach kapłanka, zwana Pytią, siadała na trójnogu i w boskiej ekstazie, w odpowiedzi na pytania błagalników, wypowiadała słowa bez związku, z których kapłan-prorok układał odpowiedzi, zwykle w formie heksametru, niekiedy zawierające błędy w metrum i stylu, które, jako pochodzące od Apollina, boga poezji, dziwiły Greków.

Władysław Kopaliński, „Słownik mitów i tradycji kultury”, hasło *Delfy*



Eugène Delacroix – Likurg radzący się pytii

# Sondaże przed I turą wyborów 2015 (z maja)

## w formie w jakiej je opublikowano

	Wybory	PP/ROBOP	IBRIS	Millward Brown	TNS Polska	Estymator	IBRIS	IBRIS
<b>Data przeprowadzenia</b>	<b>10.05</b>	<b>27.04-3.05</b>	<b>6.05</b>	<b>6-7.05</b>	<b>6-7.05</b>	<b>6-7.05</b>	<b>7.05</b>	<b>8.05</b>
<b>Liczebność próby</b>		6000	1100	3016	1000	1293	brak danych	1100
Andrzej Duda	34.76%	31.1%	29.0%	27%	27%	29%	26.8%	28.3%
Bronisław Komorowski	33.77%	39.3%	39.7%	39%	35%	40%	40.9%	36.3%
Paweł Kukiz	20.80%	12.2%	14.3%	13%	15%	16%	15.0%	17.8%
Janusz Korwin-Mikke	3.26%	5.1%	4.2%	5%	3%	4%	3.1%	3.1%
Magdalena Ogórek	2.38%	4.3%	2.4%	4%	3%	3%	2.1%	2.3%
Adam Jarubas	1.60%	3.9%	3.2%	2%	2%	4%	2.0%	1.5%
Janusz Palikot	1.42%	1.9%	1.7%	3%	1%	2%	1.6%	1.2%
Grzegorz Braun	0.83%	1.3%	0.8%	1%	1%	1%	0.9%	0.8%
Marian Kowalski	0.52%	0.6%	0.7%	0%	0%	1%	0.4%	0.7%
Jacek Wilk	0.46%	0.2%	0.0%	0%	0%	0%	0.0%	0.4%
Paweł Tanajno	0.20%	0.1%	0.1%	0%	0%	0%	0.3%	0.0%
Niezdecydowani				3.0%	13%		7.0%	7.5%
Odmowa				3.0%				
<b>Razem</b>	<b>100.0%</b>	<b>100.0%</b>	<b>96.1%</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.1%</b>	<b>99.9%</b>

Co oznacza 100%:

- w większości sondaży w 100% kryją się też niezdecydowani, a czasem odmawiający odpowiedzi;
- w niektórych sondażach procenty liczone są od liczby respondentów, którzy wskazują swojego kandydata (te „procenty” odpowiadają procentom stanowiącym potem wyniki wyborów).





# I tura – sondaże (z maja) a wybory

	Wybory	Sondaże		
	10.05	Min	Max	Średnia
Andrzej Duda	34.8%	28.7%	31.1%	29.9%
Bronisław Komorowski	33.8%	39.3%	43.9%	40.8%
Paweł Kukiz	20.8%	12.2%	19.3%	15.6%
Janusz Korwin-Mikke	3.3%	3.3%	5.3%	4.1%
Magdalena Ogórek	2.4%	2.3%	4.3%	3.2%
Adam Jarubas	1.6%	1.6%	4.0%	2.8%
Janusz Palikot	1.4%	1.1%	3.2%	1.9%
Grzegorz Braun	0.8%	0.8%	1.3%	1.0%
Marian Kowalski	0.5%	0.0%	1.0%	0.5%
Jacek Wilk	0.5%	0.0%	0.4%	0.1%
Paweł Tanajno	0.2%	0.0%	0.3%	0.1%
<b>Razem</b>	<b>100.0%</b>			

We **wszystkich** sondażach:

niedoszacowany Andrzej Duda (od 3,7 do 6,0 – średnio o 4,8 punkta procentowego)  
przeszacowany Bronisław Komorowski (od 5,5 do 10,2 – średnio o 7,0 punkta proc.)

# I tura – sondaże (z maja) a wybory

	Wybory	Sondaże		
	10.05	Min	Max	Średnia
Andrzej Duda	34.8%	28.7%	31.1%	29.9%
Bronisław Komorowski	33.8%	39.3%	43.9%	40.8%
Paweł Kukiz	20.8%	12.2%	19.3%	15.6%
Janusz Korwin-Mikke	3.3%	3.3%	5.3%	4.1%
Magdalena Ogórek	2.4%	2.3%	4.3%	3.2%
Adam Jarubas	1.6%	1.6%	4.0%	2.8%
Janusz Palikot	1.4%	1.1%	3.2%	1.9%
Grzegorz Braun	0.8%	0.8%	1.3%	1.0%
Marian Kowalski	0.5%	0.0%	1.0%	0.5%
Jacek Wilk	0.5%	0.0%	0.4%	0.1%
Paweł Tanajno	0.2%	0.0%	0.3%	0.1%
<b>Razem</b>	<b>100.0%</b>			

We **wszystkich** sondażach:

niedoszacowany Andrzej Duda (od 3,7 do 6,0 – średnio o 4,8 punkta procentowego)  
przeszacowany Bronisław Komorowski (od 5,5 do 10,2 – średnio o 7,0 punkta proc.)

Władysław Kopaliński: *Pytia [...] wypowiedała słowa bez związku, z których kapłan-prorok układał odpowiedzi, zwykle w formie heksametru, niekiedy zawierające błędy w metrum i stylu, które, jako pochodzące od Apollina, boga poezji, dziwiły Greków.*

Sondaże  
przed II turą –  
przeskalowane

	Data	Próba	Andrzej Duda	Bronisław Komorowski
Estymator	13-14.05	1107	54.0%	46.0%
Millward				
Brown	14.05	1001	52.4%	47.6%
IBRiS	14.05		52.6%	47.4%
CBOS	14-20.05	1048	52.2%	47.8%
Millward				
Brown	18.05	1256	48.4%	51.6%
Estymator	18.05	1222	52.0%	48.0%
PP/ROBOP	18-20.05	6000	51.0%	49.0%
Millward				
Brown	20.05	1257	48.9%	51.1%
IBRiS	20.05	1100	49.8%	50.2%
IBRiS	21.05	1100	49.4%	50.6%
Estymator	21.05	1341	49.5%	50.5%
TNS Polska	21.05	1000	48.9%	51.1%
Millward				
Brown	22.05	2500	50.3%	49.7%
IBRiS	22.05	1100	49.6%	50.4%
<b>Wybory</b>	<b>24.05</b>		<b>51.55%</b>	<b>48.45%</b>

# II tura – sondaże a wybory

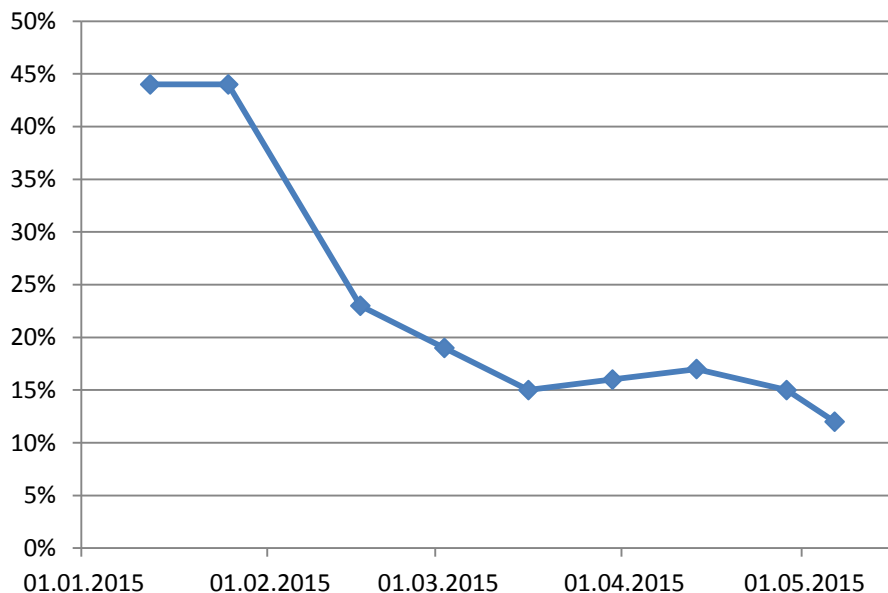
- 14 sondaży między I a II turą

	<b>Wybory</b>	<b>Sondaże</b>		
	<b>24.05</b>	<b>Min</b>	<b>Max</b>	<b>Średnia</b>
<b>Andrzej Duda</b>	<b>51.55%</b>	48.4%	54.0%	50.6%
<b>Bronisław Komorowski</b>	<b>48.45%</b>	46.0%	51.6%	49.4%

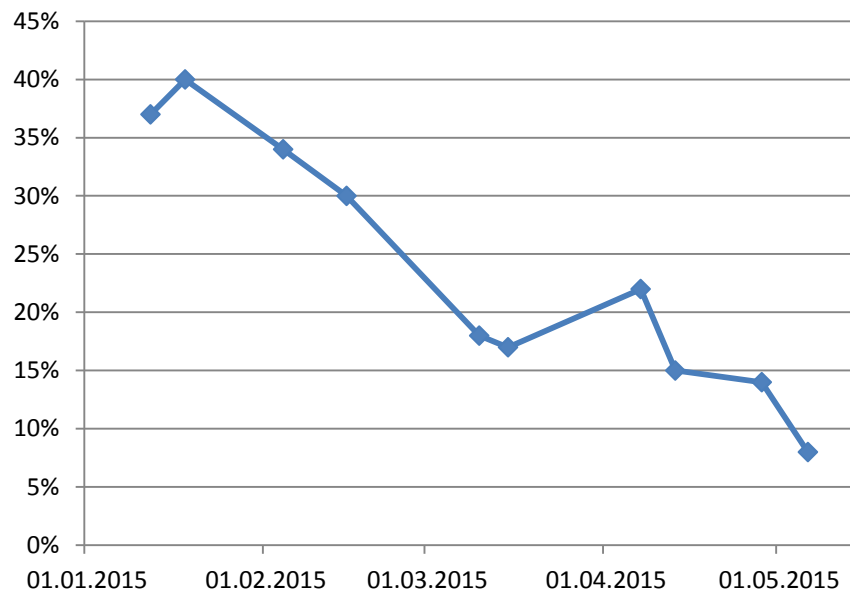
# Zanim przyłożymy nasze szkiełko i oko...

Trendy – poparcie się może zmieniać:  
**różnica Komorowski - Duda**

**Millward Brown**



**TNS Polska**



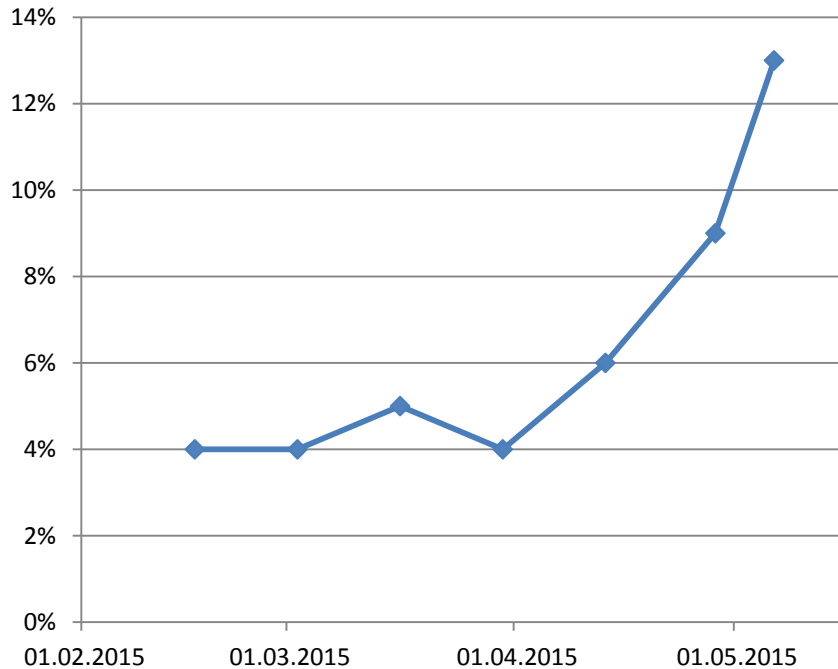
Zmiana:

około 30 punktów procentowych w ciągu 100 dni

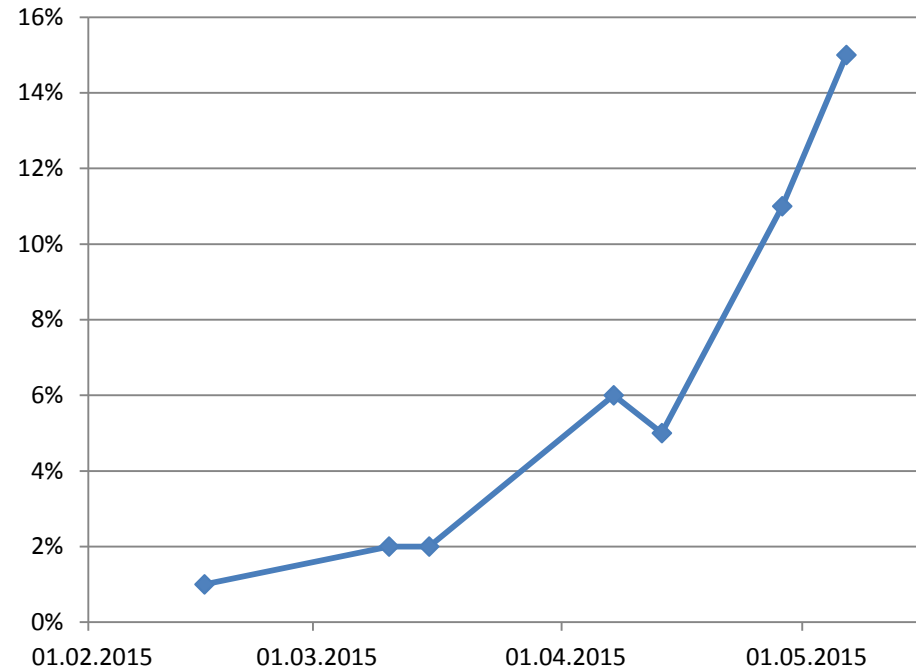
# Trendy...

## Kukiz

Millward Brown



TNS Polska



Zmiana: około 10 punktów procentowych w ciągu ostatnich 3 tygodni

# Przedmiot badań

- Fizyk badający cząstki elementarne ma przewagę intelektualną nad przedmiotem swoich badań.
- Psycholog, socjolog, sondażownie, politolog – niekoniecznie, bo badają obiekty o złożoności porównywalnej z badaczem, a nawet większej.
- Nawet w fizyce pomiar zmienia mierzony obiekt – w mechanice kwantowej (zasada nieoznaczoności Heisenberga) i nie tylko.

# Na przykład...

## Efekt Bradleya

W roku 1982 w Stanach Zjednoczonych w wyborach na gubernatora Kalifornii startowali: czarny demokrata Tom Bradley oraz biały republikanin George Deukmejian. W sondażach prowadził Tom Bradley, okazało się jednak, że przegrał wybory.



Podobny efekt zaobserwowano w niektórych innych amerykańskich wyborach, w których startowali czarny i biały kandydat.



# Wybory w Wielkiej Brytanii

- Wielka Brytania - wybory 7 V 2015

Partia	Sondaże średnia z 5 dni	Wybory	Różnica
Conservative	33.8	36.9	3.1
Labour	33.5	30.4	-3.1

Brytyjskie media były krytyczne w stosunku do swoich sondażowni:

*Lies, damn lies and election polls: Why GE2015 pundits fluffed the numbers so badly*

*Whatever you may think about the outcome of last Thursday's General Election, there is one issue on which public, politicians and pundits alike seem to be broadly united: how badly the opinion pollsters fared. They got it very wrong!*

***Kłamstwa, poważne kłamstwa i sondaże wyborcze: dlaczego eksperci wyborczy tak bardzo się pomylili.***

*Cokolwiek by nie myśleć o wynikach czwartkowych wyborów, opinia publiczna, politycy i eksperci są zgodni kwestii, jak źle wypadły sondażownie. Wypadły bardzo źle!*

- W Polsce w I turze było gorzej:

Kandydat	Sondaże średnia z 5 dni	Wybory	Różnica
Duda	29,7	34,8	-5,0
Komorowski	41,0	33,8	7,3

# Jak porównać sondaże?

Weźmy dla przykładu I turę wyborów prezydenckich w roku 2010 (wyniki po przeskalowaniu)\*:

	Wyniki wyborów 20 VI 2010	TNS Polska	GFK Polonia	IBRIS	Millward Brown	CBOS
Komorowski	41.54	40	43.2	42.5	46.3	47.2
Kaczyński	36.46	35	32.6	36.9	30.5	31.5
Napieralski	13.68	13	12.6	9.4	13.7	11.2
Korwin-Mikke	2.48	3	4.2	1.7	3.2	3.4
Pawlak	1.75	3	3.2	5.0	2.1	2.3
Olechowski	1.44	3	1.1	2.6	2.1	2.3
Lepper	1.28	3	0.0	0.9	1.1	1.1
Jurek	1.06	1	3.2	1.0	1.1	0.0
Ziętek	0.18	0	0.0	0.1	0.0	0.0
Morawiecki	0.13	0	0.0	0.0	0.0	0.0

\* Wyniki sondaży za: Michał Zieliński, Newsweek

# Który z tych sondaży jest najlepszy?

Najlepszy jest ten sondaż, który jest najbliżej wyników wyborów, czyli ten którego odległość (*metryka*) do wyników wyborów jest najmniejsza.

Wyniki sondaży i wyborów to nasza *przestrzeń metryczna* (dla 10 kandydatów mamy przestrzeń o 10 wymiarach).

Mamy wiele możliwych metryk w takiej przestrzeni.

Uwzględniliśmy te, które uważamy za najsensowniejsze.

# Metryka L\_1

- L\_1 = suma modułów błędów

$$d_m(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^n |x_k - y_k|$$

- Ta metryka zwana jest metryką taksówkową.  
Inne nazwy: metryka miejska, miasto lub Manattan.

# Metryka L\_2

- L\_2 = pierwiastek z sumy kwadratów błędów

$$d_e(\mathbf{x}, \mathbf{y}) = \sqrt{(y_1 - x_1)^2 + \dots + (y_n - x_n)^2}$$

- Ta metryka zwana jest metryką euklidesową.

# Metryka $L_\infty$

- $L_\infty = \max$  z modułów błędów

$$d_\infty(\mathbf{x}, \mathbf{y}) = \max_{k=1, \dots, n} |x_k - y_k|$$

- Ta metryka zwana jest metryką nieskończoność, maksimum, Czebyszewa, szachową.

# Kąt statystyczny

- $D_{\text{stat}}$  = arcus cosinus współczynnika Bhattacharyi

$$\Delta(x, y) = \arccos(BC(x, y))$$

gdzie  $BC(x, y) = \sum_{i=1}^N \sqrt{x_i \cdot y_i}$

- Inna nazwa: kąt Bhattacharyi
- Tę metrykę uważamy za najważniejszą i decydującą o wynikach konkursu.

# Błąd różnicy pomiędzy dwoma pierwszymi kandydatami

- $D_{12}$  = błąd różnicy pomiędzy dwoma pierwszymi kandydatami
- Ten parametr jest dodatkowy i uzupełnia wymienione wcześniej metryki.



# Porównanie sondaży z roku 2010 (I tura)

	Millward Brown	TNS Polska	IBRIS	CBOS	GFK Polonia	Wyniki wyborów 20 VI 2010
Komorowski	46.3%	40.0%	42.5%	47.2%	43.2%	41.5%
Kaczyński	30.5%	35.0%	36.9%	31.5%	32.6%	36.5%
Napieralski	13.7%	13.0%	9.4%	11.2%	12.6%	13.7%
Korwin-Mikke	3.2%	3.0%	1.7%	3.4%	4.2%	2.5%
Pawlak	2.1%	3.0%	5.0%	2.3%	3.2%	1.8%
Olechowski	2.1%	3.0%	2.6%	2.3%	1.1%	1.4%
Lepper	1.1%	3.0%	0.9%	1.1%	0.0%	1.3%
Jurek	1.1%	1.0%	1.0%	0.0%	3.2%	1.1%
Ziętek	0.0%	0.0%	0.1%	0.0%	0.0%	0.2%
Morawiecki	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
D_stat	0.0914	0.1085	0.1289	0.1444	0.1668	0.0000
L_1	0.1296	0.0986	0.1145	0.1710	0.1383	0.0000
L_2	0.0770	0.0383	0.0566	0.0825	0.0551	0.0000
L_infty	0.0599	0.0194	0.0429	0.0614	0.0389	0.0000
D_12	0.1070	0.0013	0.0051	0.1078	0.0551	0.0000

# Za chwilę porównanie sondaży z 2015 i ogłoszenie wyników konkursu o Puchar Pytii

Dziękuję,

Robert Konieczny

[robkoniec@gmail.com](mailto:robkoniec@gmail.com)