

DO

rzeczy\*



pracownia badań  
i innowacji  
społecznych



**stocznia**

\*

Jak korzystać z danych w działaniach rzeczniczych?

# Mów do rzeczy

Jak korzystać z danych w działaniach rzeczniczych?



Publikacja powstała w ramach projektu "Dane, debata, demokracja (3D)", współfinansowanego przez Szwajcarię w ramach szwajcarskiego programu współpracy z nowymi krajami członkowskimi Unii Europejskiej.

Autorzy: Jan Herbst, Jakub Rozenbaum, Jan Jakub Wygnański  
Wsparcie: Zofia Komorowska

Podręcznik powstał na podstawie materiałów przygotowanych przez autorów oraz:  
Martę Gumkowską i Agatę Miazgę ze Stowarzyszenia Klon/Jawor  
Katarzynę Starzyk i Ewę Stokłuskę z Pracowni Badań i Innowacji Społecznych „Stocznia”  
Krzysztofa Izdebskiego z Pozarządowego Centrum Dostępu do Informacji Publicznej  
Grzegorza Kowalczyka i Stanisława Mirowskiego z Millward Brown

Projekt graficzny, skład i łamanie: Studio Kotbury, [www.kotbury.pl](http://www.kotbury.pl)

Redakcja i korekta: Weronika Styś

Druk: RDS Stanisław Chłaściak, [www.reklamadrukserwis.pl](http://www.reklamadrukserwis.pl)

Nakład: 400 egz.

ISBN 978-83-62590-08-7

Warszawa 2013

Wydawca:

Fundacja Pracownia Badań i Innowacji Społecznych „Stocznia”

ul. Bracka 20B, 00-028 Warszawa

Tel: +48 22 827 01 05

[www.stocznia.org.pl](http://www.stocznia.org.pl), e-mail: [stocznia@stocznia.org.pl](mailto:stocznia@stocznia.org.pl)



Publikacja dostępna na licencji Creative Commons Uznanie autorstwa-Na tych samych warunkach 3.0 Polska.

<b>1. Po co jest ta książka?</b>	<b>7</b>
<b>2. Działania rzecznicze i wykorzystanie danych</b>	<b>10</b>
2.1 Jak rozumiemy rzecznictwo i dlaczego jest ważne?	10
2.2 Czy należy zajmować się rzecznictwem? Czy aby na pewno my, czy aby na pewno teraz?	11
2.3. Rodzaje rzecznictwa	13
2.4. Rola danych w działalności rzeczniczej	14
<b>3. Mapowanie zagadnienia</b>	<b>17</b>
3.1. Konstruowanie „drzewa problemu”	17
3.2. Konceptualizacja	20
3.3. Analiza interesariuszy	20
<b>4. Rodzaje danych</b>	<b>22</b>
4.1. Raporty i opracowania	22
4.2. Dokumenty strategiczne	25
4.3. Statystyka publiczna	26
4.4. Pozostałe dane o charakterze rejestrowym	27
4.5. Akty prawne i projekty aktów prawnych	29
4.6. Przebieg procesu legislacyjnego	30
4.7. Informacje o kluczowych aktorach	32
4.8. Monitoring mediów	33
4.9. Świadectwa	34
4.10. Wyniki badań społecznych	35
4.11. Badania własne i zlecone	36

<b>5. Zasoby danych dla działalności rzeczniczej w Polsce</b>	<b>38</b>
<b>5.1. Dane publiczne</b>	<b>38</b>
<b>5.2. Polska statystyka publiczna</b>	<b>40</b>
<b>5.3. Internetowe banki danych statystycznych</b>	<b>43</b>
5.3.1. Główny Urząd Statystyczny	44
5.3.2. EUROSTAT	49
5.3.3. Organizacja Współpracy Gospodarczej i Rozwoju (OECD)	50
5.3.4. Bank Światowy	53
5.3.5. Mojapolis.pl	54
5.3.6. Naszakasa.org.pl	60
5.3.7. Dostępne wyniki badań opinii publicznej	66
<b>6. Przeprowadzanie i zlecenie badań</b>	<b>69</b>
<b>6.1. Samodzielne przeprowadzanie badań</b>	<b>70</b>
<b>6.2. Zlecenie badań i współpraca z wykonawcą</b>	<b>72</b>
<b>7. Analiza i interpretacja danych</b>	<b>76</b>
<b>7.1. ABC statystyki – słowniczek badacza</b>	<b>76</b>
7.1.1. Agregacja	77
7.1.2. Błąd statystyczny	77
7.1.3. Indeks	79
7.1.4. Istotność statystyczna	79
7.1.5. Korelacja	81
7.1.6. Mediana (kwartyle, percentyle)	82
7.1.7. Modalna (Dominanta)	83
7.1.8. Odchylenie standardowe, wariancja	84
7.1.9. Operat losowania	84
7.1.10. Populacja	85
7.1.11. Próba	85
7.1.12. Średnia (arytmetyczna)	86
7.1.13. Ważenie	87
7.1.14. Zmienne ilościowe (interwałowe, ilorazowe) i jakościowe (nominalne, porządkowe)	87

7.1.15. Zmienne zależne/niezależne	88
<b>7.2. Rzetelność danych w rzecznictwie – na co zwracać uwagę?</b>	<b>89</b>
7.2.1. Grzech pierwszy: nieobiektywność	89
7.2.2. Grzech drugi: nieaktualność	90
7.2.3. Grzech trzeci: niekompletność, irrelewantność danych	90
7.2.4. Grzech czwarty: błędy metodologiczne	91
7.2.5. Grzech piąty: błędy techniczne i obliczeniowe	94
7.2.6. Grzech szósty: promowanie „dziwacznych” wyników	95
7.2.7. Grzech siódmy: nadinterpretacja	95
<b>8. Wizualizacja danych</b>	<b>97</b>
<b>8.1. Punkt wyjścia: jak widzimy?</b>	<b>99</b>
<b>8.2. Zasady wizualizacji</b>	<b>107</b>
8.2.1. Zasada 1. Nie pokazuj faktów, których nie ma	107
8.2.2. Zasada 2. Nie manipuluj	111
8.2.3. Zasada 3. Proste jest ładne	114
8.2.4. Zasada 4. Lepiej mniej, ale lepiej	118
8.2.5. Zasada 5. Używaj kolorów z głową	118
8.2.6. Zasada 6. Stawiaj na precyzję	120
8.2.7. Zasada 7. Opowiedz historię	122
<b>8.3. Narzędzia do wizualizacji danych</b>	<b>123</b>
8.3.1. Aplikacje dedykowane graficznej prezentacji różnych typów danych	123
8.3.2. Oprogramowanie do tworzenia różnych typów analiz i wykresów	126
8.3.3. Programy do robienia infografik	128
<b>9. Raporty rzecznicze</b>	<b>129</b>
<b>9.1. Rodzaje raportów</b>	<b>129</b>
<b>9.2. Struktura raportu</b>	<b>131</b>
<b>9.3. Formaty raportów</b>	<b>135</b>
<b>9.4. Rola raportu</b>	<b>138</b>
<b>10. Bibliografia</b>	<b>139</b>

# 1. Po co jest ta książka?

Zacznijmy od możliwie najbardziej esencjonalnego wyrażenia celu, dla którego zdecydowaliśmy się na wydanie niniejszej książki. Można go streścić jako połączenie trzech przesłanek:

- Polityki publiczne są wypadkową wielostronnych interesów i co za tym idzie, ważny jest sposób, w jaki interesy te są wyrażane.
- Organizacje pozarządowe (a przynajmniej niektóre z nich) powinny być zdolne i gotowe do wyrażania interesów (rzecznictwa) swoich członków i/lub szerzej rozumianego interesu publicznego.
- Sposób, w jaki organizacje zajmują się rzecznictwem, powinien być możliwie silnie wsparty na rzeczowych argumentach popartych dowodami – faktami.

Element pierwszy można traktować jako proste stwierdzenie faktu, natomiast drugi i szczególnie trzeci – jako rodzaje postulatów.

Naszym ambitnym celem było, aby choć trochę zmienić dotychczasową praktykę. Organizacje pozarządowe, jeśli w ogóle pojawiają się jako rzecznicy, to często sposób, w jaki to czynią (lub do którego są przez innych redukowane), sprowadza się do używania raczej emocji niż argumentów i różnych form veta zamiast propozycji rozwiązań. Takie działania są potrzebne i skuteczne jako instrumenty rzecznictwa, ale nie powinny być jedynymi. Mamy nadzieję, że lektura tej książki będzie pomocna w wykorzystywaniu różnorodnych sposobów budowania przekazu rzeczniczego.

Publikacja, którą macie Państwo w ręku, powstała w oparciu o materiały i doświadczenia zebrane w trakcie prowadzonego przez Pracownię Badań i Innowacji Społecznych „Stocznia” kursu „Rzecznictwo oparte o dowody”. Odbywał się on w ramach projektu „Dane, debata, demokracja (3D)”, który „Stocznia” realizuje w partnerstwie ze Stowarzyszeniem Klon/Jawor w ramach szwajcarskiego programu współpracy z nowymi krajami członkowskimi Unii Europejskiej.

Kurs, podobnie jak ta książka, wziął się z przekonania, że w działaniach organizacji rzeczniczych (zarówno tych funkcjonujących ogólnopolsko, jak i lokalnie) tego rodzaju podejście jest istotnie potrzebne, a żeby je prowadzić, potrzebne są określone kompetencje. Nie da się

ich nabyć wyłącznie poprzez lekturę niniejszego podręcznika, ale jest szansa, że pomoże on nieco uporządkować wiedzę w tym zakresie.

Wspomniany kurs prowadzony był od marca do maja 2013 r. w formie wykładów i warsztatów. Do uczestnictwa w nim zaprosiliśmy ponad dwudziestu przedstawicieli i przedstawicielek organizacji, które w swej pracy starają się wpływać na kształt polityk publicznych. Materiał dydaktyczny, który został zaprezentowany w czasie kursu, istotnie wykracza ponad to, co znaleźć można w niniejszym podręczniku. Wszystkich, którzy chcieliby się zapoznać z pełną wersją, namawiamy do bezpośredniego kontaktu z zespołem „Stoczni”.

Wedle naszej wiedzy jest to pierwsza tego rodzaju publikacja w Polsce. Jej ambicje nie wykraczają poza funkcje propedeutyczne. Z góry musimy uprzedzić, że wiele poruszanych tu kwestii zostało intencjonalnie opisanych w sposób dość uproszczony. W pewnym sensie niniejszy podręcznik można traktować jako rodzaj wstępu i zachęty do dalszych własnych poszukiwań.

### **Struktura podręcznika**

Staraliśmy się ułożyć podręcznik w taki sposób, żeby tworzył całość, którą najlepiej czytać zgodnie z kolejnością rozdziałów.

**Rozdział 2** to zwięzłe zarysowanie tego, czym jest i czym może być rzecznictwo, jakie są jego rodzaje oraz jaką rolę pełnią w nim dane. **W rozdziale 3** przedstawiamy pierwsze kroki, które powinniśmy wykonać, aby wykorzystanie danych było sensowne i skuteczne – jest to przede wszystkim precyzyjne określenie problemu, którym się zajmujemy i sposobu, w jaki chcemy się z nim zmierzyć.

Kolejne trzy rozdziały poświęcone są temu, jakie dane możemy wykorzystać. **Rozdział 4** to spojrzenie „z lotu ptaka” na różne rodzaje danych, które mogą być ważnym elementem stanowiącym podstawę stanowiska rzeczniczego. W **rozdziale 5** bardziej szczegółowo opisujemy, w jaki sposób pozyskiwać niektóre z nich. Skupiamy się na tych, które są ogólnodostępne w wybranych internetowych repozytoriach. W **rozdziale 6** prezentujemy kilka wskazówek, jak właściwie przygotować się do samodzielnego gromadzenia danych poprzez własne badania w sytuacji, gdy dane zastane okazują się niewystarczające.

**Rozdziały 7, 8 i 9** dotyczą tego, co z można zrobić z już zgromadzonymi danymi. **Rozdział 7** to wprowadzenie do analizy i interpretacji danych. Znajdziemy tam słowniczek z istotnymi



pojęciami – przede wszystkim z zakresu podstaw statystyki oraz najważniejsze zasady, które pozwolą zapewnić rzetelność wykorzystanych danych. **Rozdział 8** to część poświęcona wizualizacji danych – zasadom, zgodnie z którymi należy tworzyć wizualizacje, błędom, których trzeba unikać i narzędziom, z których warto korzystać. Natomiast **rozdział 9** to opis zwieńczenia prac analitycznych, a mianowicie konstruowania tzw. raportów rzeczniczych. Staraliśmy się w nim także przedstawić konkretne przykłady raportów, które z różnych powodów wydały nam się ciekawe i inspirujące.

Ostatnia część podręcznika (**rozdział 10**) to bibliografia – wybór publikacji i źródeł, które pozwalają na pogłębienie tematyki każdego z rozdziałów.

## 2. Działania rzecznicze i wykorzystanie danych

### 2.1. Jak rozumiemy rzecznictwo i dlaczego jest ważne?

Rozsądnie byłoby zacząć od odpowiedzi na pytania, na ile w ogóle ważne jest rzecznictwo w działaniach organizacji i na ile mają one realny wpływ na tworzenie polityk publicznych. Odpowiadając na pierwsze z pytań, warto zdać sobie sprawę z tego, że rzecznictwo to tylko jedna z funkcji, jakie pełnią organizacje. Nie wszystkie zresztą powstały w tym celu i powinny się nim zajmować. Znacznie częstsze powody ich powstawania i działania to dostarczanie konkretnych usług i zaspokajanie potrzeb członków (szczególnie tam, gdzie zawodzi państwo lub rynek), tworzenie przestrzeni dla wspólnych pasji i zainteresowań czy pielęgnowanie różnych form altruizmu.

Wszystko to prawda, ale organizacje pełnią także istotne funkcje natury politycznej, przyczyniając się do swoiście rozumianej równowagi władzy – w szczególności poprzez „patrzenie jej na ręce” oraz wyrażanie interesów poszczególnych grup społecznych, czyli właśnie rzecznictwo. Dla niektórych typów organizacji rzecznictwo jest właściwie działaniem ciągłym, dla innych epizodem związanym z konkretną „sprawą do załatwienia”. Według ostatniego badania Stowarzyszenia Klon/Jawor rzecznictwem rozumianym jako wpływanie na rozwiązania systemowe zajmuje się ok. 7% organizacji. Przy okazji warto odnotować, że ich odsetek w ciągu ostatnich 10 lat spadł prawie dwukrotnie, co nie jest niestety dobrą wiadomością.

Myśląc o rzeczniczej powinności organizacji, dotykamy niezwykle istotnej kwestii, jaką jest rola tzw. instytucji pośredniczących w funkcjonowaniu demokracji. W tej sprawie istnieje szereg opinii, ale dla uproszczenia przywołajmy tylko dwie – skrajnie przeciwstawne.

Wedle pierwszego stanowiska istnienie organizacji jest absolutnie niezbędne dla dobrego funkcjonowania demokracji. Czasem mówi się o tzw. demokracji asocjacyjnej, którą opisuje znana klasyczna książka A. Tocqueville’a „O demokracji w Ameryce”. Fundamentem tej konstrukcji jest przekonanie, że to właśnie działania organizacji (w szczególności stowarzyszeń) tworzą rodzaj „immunologii” dla ewentualnych patologii władzy. Bez ich działania indywidualni obywatele będą zagrożeni zostawieni „sam na sam” z wszechpotężnym państwem,

---

1 Raport z badania Stowarzyszenia Klon/Jawor „Kondycja sektora organizacji pozarządowych w Polsce 2012”

które łatwo może wtedy zamienić się w opresora. Można też jednak spotkać pogląd zgoła przeciwny, zgodnie z którym różnego rodzaju stowarzyszenia, korporacje zawodowe itp. psują i w pewnym sensie „przechwytyją” demokrację.

Jakkolwiek paradoksalnie to zabrzmiałoby, obydwa te poglądy mogą być prawdziwe. W istocie wiele zależy od jakości owych instytucji pośredniczących. Sprawa jest niezwykle delikatna. Dobrym prawem organizacji jest wyrażanie interesów swoich członków, ale trzeba też pamiętać, że interesy poszczególnych grup nie zawsze i nie automatycznie zbieżne są z interesem publicznym. Oczywiście są też organizacje, które z definicji poszukują dobra publicznego i starają się je chronić (stąd też znane w wielu krajach rozróżnienie na organizacje pożytku publicznego oraz organizacje wzajemnościowe).

Żeby było jasne, nie twierdzimy, że jedne organizacje są dobre, a inne złe. Przywołane rozróżnienie jest jednak wciąż dość istotne i prowadząc działania rzecznicze, warto o nim pamiętać. To zresztą jeszcze jeden argument za tym, aby nasze działania były dobrze i uczciwie uzasadnione i abyśmy z szacunkiem podchodzili do poglądów wyrażanych przez innych, nawet jeśli są odmienne od naszych.

Ten nieco dydaktyczny ton jest chyba potrzebny, żeby równoważyć przesadne czasem poczucie moralnej wyższości, jakie wyrażają organizacje pozarządowe w relacjach z władzami publicznymi. Warto pamiętać, że pełnienie przez organizacje niezwykle ważnej roli „strażników” w stosunku do władz nie zwalnia, ale wręcz podkreśla wagę swoiście rozumianej samodyscypliny. Wisi nad nimi ciągle antyczne pytanie: „kto będzie strażnikiem strażników”? W tym kontekście warto też przypomnieć w charakterze przestrogi, jak mierne sukcesy ma środowisko pozarządowe we wdrożeniu przyjętej ponad 10 lat temu Karty Zasad Organizacji. Nie brakuje też ostatnio przykładów, gdzie organizacje raczej rujną niż budują swoją reputację jako rzeczników poprzez nadużycia lub złe, naciągane czy manipulacyjne użycie danych.

## **2.2. Czy należy zajmować się rzecznictwem?**

### **Czy aby na pewno my, czy aby na pewno teraz?**

Jakkolwiek dziwnie to zabrzmiałoby w kontekście faktu, że po tę książkę sięgną zapewne zainteresowani rzecznictwem – warto dobrze przeanalizować, czy i w jaki sposób w ogóle warto angażować się w działania rzecznicze. W tym celu dobrze uprzednio odpowiedzieć sobie na kilka pytań:

- Jaka jest pozycja i kompetencje naszej organizacji – czy pozwalają nam one na skuteczne podjęcie działań rzeczniczych?
- W jakim otoczeniu działamy, czy potrafimy zmobilizować innych do działania?
- Czy mamy dostatecznie ugruntowaną reputację i wiarygodność, aby być skutecznym „ambasadorem” sprawy, o którą zabiegamy?
- Czy istotnie to my powinniśmy się tym zajmować?
- Czy mamy poparcie tych, w których imieniu zamierzamy się wypowiadać? Czy mamy ich „mandat”, czy nie ryzykujemy zarzutu pewnego rodzaju uzurpacji?
- Czy nasze motywacje są autentyczne i czyste – innymi słowy, czy idzie nam o załatwienie ważnych spraw, czy może swoiście rozumianą autopromocję?
- Czy dobrze rozumiemy konsekwencje rozwiązań, które sami proponujemy? W szczególności – czy mamy pewność, że nie będą to zmiany na gorsze?
- Czy umiemy zarządzać ryzykiem i gotowi jesteście odpierać różnego rodzaju ataki, które pojawiają się wtedy, gdy zaczynamy głosić niekoniecznie popularne poglądy?
- Co szczególnie ważne z punktu widzenia tej książki – czy mamy przekonujące dla nas samych i dla innych dowody na poparcie formułowanych przez nas poglądów?
- Wreszcie – czy dobrze przeanalizowaliśmy, jakich spraw i działań (być może istotniejszych z punktu widzenia misji organizacji) nie będziemy mogli podjąć poświęcając energię na działania rzecznicze?

Powyższe kłopotliwe nieco pytania nie mają rujnować motywacji do działań rzeczniczych, ale spowodować, że zabierzemy się za nie odpowiedzialnie i z nadzieją na sukces.

W formie pytań staraliśmy się uświadomić czytelnikowi, jakie korzyści, ale i koszty wynikają z podjęcia przez organizację działań rzeczniczych. Spójrzmy teraz na nie z punktu widzenia tych, do których są kierowane – w szczególności szeroko rozumianej kategorii władz i decydentów. Zaczniemy od tego, że dla wielu z nich korzyści wynikające z działań rzeczniczych nie są wcale oczywiste. Rzecznictwo – szczególnie prowadzone źle – może być łatwo uznane za komplikujące proces decyzyjny, stronnictwo, wręcz podejrzane i kojarzące się z nie zawsze godnymi pochwałą formami lobbingu. Często jest kłopotliwe, bo wskazuje na być może niedostateczną wiedzę czy wrażliwość samych decydentów. Może być też traktowane (w przypadku władz wybieralnych) jako podważanie ich demokratycznego mandatu.

Wiele z tych obaw jest na wyrost, a część z nich zasadza się na nieporozumieniach. Bywa, że są one po prostu swoistą racjonalizacją niechęci do wchodzenia w relacje z partnerami zewnętrznymi. Co mogłoby ich przekonać do zmiany stanowiska? Ostatnie lata dostarczają

nowych argumentów w tej sprawie. Proces rządzenia przybiera coraz bardziej interaktywne formy. Coraz bardziej trafną metaforą rządzenia jest raczej sieć niż hierarchia. Sieć zaś oznacza funkcjonowanie w wielostronnych relacjach i konieczność brania pod uwagę różnych interesów i punktów widzenia.

Z jednej strony to dobrze, z drugiej stanowi to poważne wyzwanie, bo tworzenie spójnej polityki w tej sytuacji jest prawdziwym wyzwaniem. Wielu uczestników tak rozumianej sieci ma skłonność do widzenia spraw w wąskiej perspektywie. Postaci tych jest coraz więcej, nie są to już przedstawiciele tradycyjnych tzw. partnerów społecznych (pracodawcy i związki zawodowe), ale także organizacje (czasem określa się to jako przesunięcie akcentów z dialogu społecznego na dialog obywatelski). Te ostatnie są na ogół z różnych powodów na gorszych pozycjach niż wymienione wcześniej grupy. Warunki ich dostępu do procesu konsultacyjnego nie są na ogół ugruntowane w przepisach. Często muszą się one wręcz dobijać o prawo do wyrażania swoich poglądów. Organizacje, a nawet ich federacje, często nie posiadają też etatowych ekspertów, lobbystów i w ogóle osób, które mogłyby się poświęcić np. monitorowaniu bieżących propozycji legislacyjnych.

Warto też pamiętać, że obecnie coraz częściej w sprawach publicznych wypowiadają się bezpośrednio poszczególni obywatele lub tworzone przez nich *ad hoc* grupy – niekiedy zwoływane wyłącznie w celu załatwienia (częściej zablokowania) konkretnej decyzji. Można powiedzieć, że demokracja nabiera coraz częściej osobistego i nieomal codziennego charakteru.

### **2.3. Rodzaje rzecznictwa**

Na potrzeby tej książki użyteczne będzie uporządkowanie różnych rodzajów rzecznictwa. Zatem jedne odnoszą się do indywidualnych osób (i działania w ich imieniu). Inne polegają na reprezentowaniu grup czy kategorii osób. Wreszcie są też takie, które dotyczą szerszej rozumianych dóbr publicznych. Typy te nieraz w praktyce łączą się i przenikają, często wynikają jedne z drugich. Dla nas, z punktu widzenia tej publikacji, najważniejsze będą te formy rzecznictwa, które wykraczają poza rzecznictwo indywidualne. Czasem nazywane są one wręcz rzecznictwem politycznym (systemowym) w odróżnieniu od rzecznictwa osobistego. Przymiotnik „polityczne” w tym kontekście należy rozumieć rzecz jasna w sposób szerszy niż politykę partyjną – chodzi raczej o odniesienie do polityk publicznych i działania władz publicznych.

Różne mogą być też cele rzecznictwa. Czasem chodzi o wywołanie zmiany, innym razem wręcz przeciwnie – o to, aby ją powstrzymać (jeśli ma to być zmiana na gorsze z punktu widzenia określonej grupy). Celem działań rzeczniczych mogą być m.in.: zmiany regulacji prawnych – wprowadzenie nowych ustaw czy przepisów wykonawczych; zmiana interpretacji już wydanych przepisów; realizacja/egzekucja praw; odmienna alokacja środków; zmiana w funkcjonowaniu instytucji publicznych.

Różne mogą też być elementy (instrumentarium) prowadzenia działań rzeczniczych. Skuteczne ich prowadzenie wymaga najczęściej połączenia kilku z tych narzędzi. Niejako z definicji rzecznictwo oznacza przede wszystkim wyrażanie i docieranie z określoną opinią do różnych grup. Zazwyczaj chodzi o decydentów, ale często może także zależeć nam na dotarciu do szerszej rozumianej opinii publicznej (zmiana nastawienia tych pierwszych może być wywołana poprzez nacisk tych drugich). Z tego też powodu elementem działań rzeczniczych są często różnego rodzaju kampanie społeczne, a ich istotnym elementem (istotną częścią przekazu) mogą być właśnie rzetelne i dobrze zaprezentowane dane. Skuteczne działanie rzecznicze wymaga też umiejętności budowania szerszych koalicji zdolnych do kolektywnego działania – często jest to wręcz warunek sukcesu. Do repertuaru działań rzeczniczych należą też czasem działania lobbingowe (rozumiane w sposób tradycyjny – oznaczający bezpośrednie, także nieformalne kontakty z decydentami). Istotnym elementem wpływu na sprawy publiczne w przypadku organizacji jest uczestnictwo w różnego rodzaju ciałach kolegialnych opiniodawczych i doradczych – jest ich obecnie coraz więcej (np. Rady Pożytku Publicznego różnych szczebli). Nie można wykluczać z repertuaru działań rzeczniczych także i takich, których istotą jest protest i niezgoda, które mogą przybierać formę otwartego konfliktu i to nie zawsze prowadzonego w parlamentarnym stylu. Forma ta powinna być traktowana jako ostateczność, choć w praktyce z różnych powodów uruchamiana jest ona dość często. Najważniejsze, że bez względu na to, które z powyższych metod zastosujemy, w każdej z nich istotnym elementem mogą i powinny być dobrze udokumentowane i zaprezentowane fakty i dowody.

## **2.4. Rola danych w działalności rzeczniczej**

Założeniem tej książki jest to, że rzecznictwo oparte o dowody (dane) jest racjonalnym i skutecznym sposobem postępowania. Innymi słowy, że w procesie formułowania polityk publicznych ważne są rzeczowe argumenty i udokumentowane fakty dotyczące samej natury problemu. Praktyka wygląda czasem nieco inaczej i górę biorą zgoła inne argumenty i powody. Ich lista jest dość długa: koniunkturalizm i oportunizm, klientyzm, upartyjnienie, osobiste

koneksje, chęć uniknięcia ryzyka, źle rozumiany formalizm i wiele innych. Wszystkie one jednak nie powinny nas zniechęcać do konsekwentnych zabiegów na rzecz merytorycznej jakości debaty, której podstawą powinny być dane w możliwie szerokim zakresie. Do czego mogą się one przydać i dlaczego warto je mieć?

Zacząć trzeba od tego najważniejszego – dane ugruntowują i uzasadniają nasze postulaty i nadają im treść. To, że potrafimy je gromadzić i roztropnie ich używać, tworzy i podtrzymuje autorytet organizacji i wskazuje na to, że w debacie powinna być traktowana jako ekspert i tym samym – ważny jej uczestnik. W oparciu o dane można nie tylko budować bardziej wiarygodne diagnozy, ale także szukać i modelować propozycje rozwiązań problemów. Dobrze zaprezentowane dane (czy raczej dokonane na ich podstawie syntezy) tworzą często niezbędną – z punktu widzenia skuteczności działań komunikacyjnych – ilustrację skali problemów i metod, którymi chcemy je rozwiązać. Z tego samego powodu dane są jedną z najczęściej przyswajanych i cytowanych przez media informacji. Często obiektywne dane są podstawowym mechanizmem walki ze stereotypami utrudniającymi znalezienie porozumienia.

Zaletą posiadania danych jest też to, że często tylko w oparciu o nie można sformułować wiarygodne kalkulacje kosztów, które towarzyszą występowaniu poszczególnych problemów i związanych z nimi zaniechań. Dane pozwalają też tworzyć wiarygodne przewidywania dotyczące kosztów działań naprawczych. Bez tego rodzaju wyliczeń formułowane przez nas diagnozy oraz propozycje rozwiązań są nierzadko ignorowane. Wreszcie dane pozwalają nadać sprawom właściwe proporcje i skalę, a tym samym – odpowiednią wagę.

W działaniach rzeczniczych nie chodzi o jakiegokolwiek dane, ale o dane wiarygodne. Rozumieć trzeba przez to wiarygodność źródeł i zastosowanych metod. Dane, którymi operujemy, powinny być relewantne do tego, o czym mówimy, a także aktualne i osadzone w kontekście, którego problem dotyczy. Prezentując dane, powinniśmy liczyć się z tym, że możemy zostać zapytani o ich źródło i przyjęte założenia metodologiczne. Jeśli dane mają być instrumentem zmiany, muszą koncentrować się na tym, co rzeczywiście istotne. Trzeba wystrzegać się powszechnej także wśród organizacji manieri prezentowania olbrzymiej ilości danych pozbawionych interpretacji i odpowiedzi na pytanie, co właściwie z nich wynika. Raz jeszcze warto przypomnieć, że dane są jedynie składnikiem diagnoz, a nie ich substytutem. Musimy być zdolni do wyciągania nie tylko wniosków, ale też konkluzji i rekomendacji do działania (innych rzecz jasna niż te, że potrzebne są kolejne badania). W wyciąganiu wniosków musimy zachować intelektualny rygor, wystrzegać się życzeniowego myślenia, prowadzenia badań

z wynikiem zamierzonym z góry i wszelkich innych form manipulacji, na których na ogół prędzej czy później zostaniemy przyłapani. Wtedy nie tylko nie pomożemy, ale wręcz zaszkodzimy sprawie, o którą zabiegamy. Musimy też umieć dopasować prezentowane dane do typu audytorium, z jakim mamy do czynienia. Należy wystrzegać się żargonu, pseudo-naukowości i źle pojętego akademizmu. W praktyce oznacza to, że czasem trzeba dysponować więcej niż jedną formą przekazu zgromadzonego materiału. Często konieczne jest dokonanie skrótów, wyciągnięcie esencji lub wręcz narysowanie tego, co najważniejsze w naszym przekazie.



### 3. Mapowanie zagadnienia

W podejściu, które proponujemy, dane – rozumiane tu bardzo szeroko – odgrywają kluczową rolę. Nie powinny one stanowić koła ratunkowego, które staramy się znaleźć ad hoc w sytuacji, w której ktoś spyta się nas: „a skąd Pani/Pan to wie?”. Wykorzystanie danych może i powinno być uporządkowanym procesem, przemyślanym pod kątem celu, jaki chcemy realizować w naszej działalności rzeczniczej. Proces ten trzeba zaplanować, zaczynając od etapu „mapowania zagadnienia” (ang. issue mapping).

To pojęcie można różnie rozumieć. Czasem wykorzystuje się je do określenia bardzo konkretnych, technicznych czynności, często wspomaganych programami komputerowymi. Tutaj chcielibyśmy potraktować je szerzej, jako dogłębne przeanalizowanie zagadnienia będącego przedmiotem naszej działalności w jego różnych aspektach. Celem powinno być ustalenie, jakich danych poszukujemy, gdzie powinniśmy ich szukać i do czego chcemy je wykorzystywać.

#### 3.1. Konstruowanie „drzewa problemu”

Punktem wyjścia takiego zabiegu może być rozłożenie problemu na czynniki pierwsze pod kątem przyczynowo-skutkowym, co nazywane jest często tworzeniem „drzewa problemu”. Problem, którym się zajmujemy, nie jest bowiem na ogół wyizolowanym zjawiskiem społecznym (objawem), lecz znajduje się w sieci powiązań różnych zjawisk i przyczyn.

Posłużmy się tutaj przykładem problemu **niskiej dzietności w Polsce**.

Żeby dobrze go zrozumieć, niezbędne jest **zdiagnozowanie przyczyn**. Być może, żeby wpłynąć na zmianę sytuacji demograficznej w Polsce, której istotnym i niepokojącym elementem jest obecnie bardzo niska dzietność, trzeba przede wszystkim przeanalizować powody, dla których rodzice podejmują decyzje o założeniu i powiększeniu rodziny (na przykład dostępność mieszkań, poczucie bezpieczeństwa zatrudnienia etc.). Takich czynników może być, rzecz jasna, więcej. W większości przypadków nie można się zatrzymywać na najbardziej bezpośrednich przyczynach, ponieważ często rzeczywiste źródło problemu leży dość głęboko.

Aby ukazać pełne znaczenie zjawiska, trzeba również spojrzeć na jego **skutki**. W naszym przykładzie możemy za podstawowy problem uznać sam fakt zmniejszającej się dzietności, ale warto jednak pomyśleć również o tym, jakie inne przemiany społeczne to za sobą pociąga

## przyczyny

- niski dostęp do mieszkań (w stosunku do zarobków)
- brak dostępu do opieki nad dziećmi
- zmiany w modelu rodziny (niedostępni dziadkowie)
- dziecko jako towar luksusowy
- niestabilna sytuacja ekonomiczna ludzi
- wysoki koszt społeczny (jedna wypada z życia społecznego)
- brak polityki przywracania matek na rynek pracy
- zmiana stylu życia - różnorodność ról - podejście do rodzicielstwa
- świadome rodzicielstwo (możliwość podjęcia decyzji)
  - język zachęt ekonomicznych
- brak wsparcia dla rodzin wielodzietnych
- migracja - brak sieci wsparcia w środowisku lokalnym
- brak wsparcia dla rodzin z dziećmi z chorobami
- polityka reprodukcyjna w Polsce - brak podmiotowości kobiet

## konsekwencje

- starzenie się społeczeństwa
- więcej osób wymagających opieki i mniej mogących jej udzielić
- zmniejszająca się liczba osób w wieku produkcyjnym
- mniej osób płacących podatki i składki
- spowolnienie gospodarki
- dla jednostki - brak doświadczenia rodzicielstwa
- frustracja osób, które nie mogą sobie pozwolić na dziecko

## Niska dzietność w Polsce

Obraz 3.1. Roboczy schemat przyczyn i skutków niskiej dzietności w Polsce wykonany w ramach ćwiczenia przez uczestników kursu „Rzecznictwo oparte o dowody”.

– może się okazać, że to one są dużo ważniejsze. Podobnie jak w przypadku przyczyn, warto patrzeć nie tylko na bezpośrednie skutki, ale i na te dalekosiężne, bo czasem to dopiero one nadają rzeczywistą wagę zagadnieniu, którym się zajmujemy.

Wiedza o przyczynach zjawiska pozwala wskazać, gdzie powinno się podjąć interwencję, żeby zmierzać w stronę rozwiązania problemu. Z kolei znajomość skutków uświadamia, jakie są koszty indywidualne, społeczne, ekonomiczne zastanej sytuacji, a także – co może się zmienić, jeśli podejmiemy odpowiednie działania. W niektórych sytuacjach możliwe jest też określenie odroczonego w czasie, kumulujących się kosztów czy szerszej konsekwencji danego stanu rzeczy. Często dopiero taki „rachunek ciągłony” pozwala przeprowadzić i określić wysokość realnych kosztów zaniechania i pozostawienia spraw własnemu biegowi. Może to być trafnym argumentem w debacie. Szczególnie wtedy, gdy chodzi o uruchomienie działań, których horyzont wykracza poza rok budżetowy czy kadencję wyborczą.

Często już ten właśnie etap tworzenia „drzewa problemu” wymaga oparcia się o dane. W tym wypadku z reguły posiłkujemy się danymi najbardziej podstawowymi, już przez kogoś opracowanymi, albo wykorzystujemy doświadczenie i dobrą znajomość tematu – własną i innych organizacji.

Tworzenie „drzewa problemu” to jeden z tych etapów, w których współpraca z innymi osobami czy organizacjami jest nieodzowna, by osiągnąć dobre efekty. Nawet jeśli specjalizujemy się w danym zagadnieniu, patrzymy na nie z określonej perspektywy i zwracamy uwagę tylko na część jego aspektów. Tylko wspólna dyskusja w gronie różnych osób i instytucji pozwala na stworzenie w miarę pełnego obrazu problemu. Do takiej rozmowy warto zaprosić nie tylko inne organizacje – często wartościowy jest głos przedstawicieli administracji lub instytucji publicznych, osób ze świata nauki, a w niektórych przypadkach również przedsiębiorców. **[patrz obraz 3.1. str.18]**

Warto zwrócić uwagę, że analizę przyczyn i skutków można zastosować w równym stopniu do problemów o skali krajowej czy globalnej, jak i lokalnych bolączek konkretnej gminy czy wsi. Ilustracją może być tutaj np. utrudniony dostęp do informacji publicznej w gminie X. Poza różnymi lokalnymi zjawiskami, tego typu problem ma na pewno wiele uwarunkowań systemowych, o których nie można absolutnie zapominać, jeśli chcemy go rozwiązać. Ale również lokalne przyczyny często układają się we wzory powtarzające się w różnych innych gminach, co daje nam większe możliwości zdobycia wiedzy o istocie problemu.

## 3.2. Konceptualizacja

„Drzewo problemów” pozwala nam w sposób uporządkowany rozpoznać, jakimi szczegółowymi kwestiami powinniśmy się zajmować, jeśli naszym celem jest rozwiązanie danego problemu.

Następnym krokiem jest sformułowanie w oparciu o to pytań, na które chcemy znaleźć odpowiedzi, korzystając z danych. Wracając do naszych przykładów, takimi pytaniami mogą być np. „Ile jest osób, które nie decydują się na drugie dziecko i jakie są tego przyczyny?” albo „Jakie są reakcje osób, które uzyskały odpowiedź odmowną na wniosek o udostępnienie informacji publicznej?”

W badaniach społecznych ten etap określany jest jako „konceptualizacja” i jest niezbędny do tego, żeby badanie przyniosło zamierzone efekty. Poszukiwanie danych także jest swego rodzaju badaniem (tzw. desk research), nawet jeśli ogranicza się do ogólnodostępnych źródeł. W związku z tym ono również wymaga od nas przejścia przez etap konceptualizacji. W tym momencie powinniśmy zejść z poziomu ogólnych problemów, takich jak nasze przykładowe: niska dzietność czy utrudniony dostęp do informacji publicznej, i rozłożyć je na mniejsze części składowe, które precyzyjnie zdefiniujemy. W przypadku danych ilościowych możemy wręcz mówić o konstruowaniu wskaźników, którymi chcemy mierzyć poziom danego zjawiska. Przykładowy problem dostępu do informacji publicznej zdołamy przełożyć na takie konkretne elementy jak np. „odsetek odpowiedzi odmownych na wnioski o udostępnienie informacji publicznej” czy (odpowiednio zdefiniowany) „stosunek urzędników do udostępniania informacji”.

Poważne potraktowanie tego elementu przygotowań do gromadzenia danych bardzo ułatwi nam dalsze kroki. Sformułowanie konkretnych pytań pozwoli stwierdzić już na samym początku, jakiego rodzaju danych poszukujemy **[patrz rozdział 4. „Rodzaje i źródła danych”, str. 22]**, czego mają dotyczyć i gdzie oraz w jaki sposób ich szukać.

## 3.3. Analiza interesariuszy

Kolejnym elementem mapowania zagadnienia rzeczniczego jest tzw. analiza interesariuszy danego problemu oraz analiza odbiorców naszych działań rzeczniczych. Są to dwie podobne, ale jednak nie tożsame kategorie.

Po pierwsze, możemy zastanawiać się nad tym, kto w kwestii, którą się zajmujemy, odgrywa ważną rolę. Jakie grupy, jakie konkretne osoby z jednej strony – mają dużo do powiedzenia, a z drugiej – ponosić będą konsekwencje danego działania (lub zaniechania). W niektórych przypadkach ten obszar wiedzy o danym zagadnieniu jest kluczowy i to w tym kierunku będą zmierzać poszukiwania danych.

Po drugie, od samego początku musimy myśleć o tym, do kogo planujemy adresować nasze działania rzecznicze. Czy chcemy zwracać się przede wszystkim do władz? Państwowych, samorządowych, a może europejskich – bo tylko na tym poziomie da się rozwiązać dany problem? Czy może raczej ważniejsze jest, żeby usłyszały nas media, które nagłośnią sprawę w taki sposób, że stanie się ona przedmiotem debaty publicznej? Możemy też skierować swoje działania przede wszystkim do tych, których rzecznikami jesteśmy – np. do bezdomnych, samotnych matek, rodziców dzieci na terenach wiejskich itp. A być może najważniejszym odbiorcą naszego przekazu będzie sponsor, który pomoże zrealizować nasz pomysł, dając nadzieję na rozwiązanie problemu?

Nie ma tu jednej dobrej odpowiedzi. W zależności od sytuacji każda z tych opcji może być uznana za najlepszą. Nie wykluczają się one też wzajemnie i czasami wskazane jest kierowanie się z różnymi aspektami danego problemu do różnych odbiorców.

Jednak najważniejsze jest, by wiedzieć od początku, na co się decydujemy. Od tego bowiem zależy, na jakie dane będziemy zwracać szczególną uwagę.

Podsumowując, żeby efektywnie wykorzystać nasz czas i energię, trzeba mieć – w głowie lub na papierze – strategię wykorzystania danych w naszej działalności. Powinniśmy:

- zdefiniować problem, jego przyczyny i skutki;
- wiedzieć, do kogo chcemy się zwracać z naszym przekazem/naszymi działaniami;
- wstępnie rozpoznać, jakie dane mamy do dyspozycji w temacie, którym się zajmujemy;
- zorientować się, jakimi zasobami (finansowymi, czasowymi itp.) dysponujemy.

Dopiero wtedy możemy zabrać się na poważnie za gromadzenie, analizę i wykorzystanie danych. Pamiętajmy przy tym o celu, do jakiego mają one służyć – z reguły będzie to dostarczenie solidnej i wiarygodnej podstawy dla naszego stanowiska rzeczniczego.

## 4. Rodzaje danych

Kiedy mówimy o danych, z reguły mamy na myśli różne liczby: odsetki, liczebności, stopy, współczynniki itp. Jednak dane możemy rozumieć również szerzej. Dane mają nam coś powiedzieć o zagadnieniu, którym się zajmujemy.

Czasami głos jednej osoby może przekazać więcej niż sto liczb [patrz podrozdział 4.9 „Świadectwa”, str. 34] – zwłaszcza, że ostatecznie za większością liczb, którymi się posługujemy, stoją realni ludzie. Również treść różnego rodzaju dokumentów może być dla nas kluczową informacją [patrz podrozdział 4.2 „Dokumenty strategiczne”, str. 25]. Krótko mówiąc, dane to nic innego jak zapisana w określonej formie informacja o interesujących nas faktach, zjawiskach lub procesach.

Warto więc spojrzeć na to, z jakich rodzajów danych możemy skorzystać. Poniższa typologia nie jest ani wyczerpująca, ani w pełni rozłączna – ma jednak służyć uporządkowaniu myślenia o tym, do jakiego typu informacji możemy sięgnąć zależnie od założonego celu.

### 4.1. Raporty i opracowania

#### Krótką charakterystyka

Podstawowym elementem naszego „morza danych” jest to, co na interesujący nas temat już zostało napisane. Istniejące opracowania mogą stanowić samoistne źródło danych, ale mogą też wyznaczać punkt wyjścia do dalszych poszukiwań, wskazywać na inne potencjalne źródła informacji.

#### Kiedy z nich korzystać?

Do raportów powinniśmy sięgnąć zawsze, aby sprawdzić, czy to, czego chcemy się dowiedzieć, nie zostało już opisane.

Powoływanie się na raporty wydane przez rozpoznawalne, posiadające autorytet instytucje może nam pomóc w sytuacji ryzyka, że stanowisko organizacji pozarządowej będzie dezawuowane jako stroniczne lub nie dość poważne. Z kolei niezajomość uznanych opracowań i raportów może narazić nas na zarzut ignorancji, śmieszność czy utratę mandatu do zabierania głosu w danej sprawie.

## Gdzie je znaleźć?

Poza instytucjami zajmującymi się określonymi dziedzinami, istnieją również takie, które niejako zawodowo zajmują się tworzeniem raportów na różne tematy, np. think tanki (swoiste „rezerwuary/zbiorniki idei”).

W Polsce takie organizacje są jednak wciąż dużo mniej rozwinięte niż np. w Niemczech, Anglii czy w Stanach Zjednoczonych, gdzie regularnie proponują one nowe rozwiązania w różnych obszarach polityk publicznych. Instytucje takie często reprezentują określony kierunek polityczny czy specyficzną, nazwijmy to, wrażliwość w sprawach publicznych. Nierzadko działają np. przy partiach politycznych (np. tak jest w Niemczech, gdzie działają fundacje wprost finansowane ze środków publicznych działające przy partiach politycznych, np. Fundacje im. Friedricha Eberta, im. Konrada Adenauera, im. Heinricha Bölla). Działają też różnego rodzaju instytucje branżowe wyspecjalizowane w jakimś zagadnieniu, np. kwestiach środowiska naturalnego, finansów publicznych, prawa i warunków pracy etc.

W Europie Środkowo-Wschodniej od początku transformacji podejmowano liczne starania, aby stworzyć warunki dla działania tego rodzaju instytucji. Powstało nawet ich porozumienie Policy Association for an Open Society (<http://pasos.org>).

---

Do rozpoznawalnych polskich think tanków należą m.in.:

Centrum Analiz Społeczno-Ekonomicznych CASE (<http://www.case-research.eu/pl>)

Centrum im. Adama Smitha (<http://www.smith.org.pl/>)

Demos Europa ([www.demoseuropa.eu](http://www.demoseuropa.eu))

Forum Obywatelskiego Rozwoju ([www.for.org.pl](http://www.for.org.pl))

Instytut Badań nad Gospodarką Rynkową (<http://www.ibngr.pl/>)

Instytut Badań Strukturalnych ([www.ibs.org.pl](http://www.ibs.org.pl))

Instytut Obywatelski ([www.instytutobywatelski.pl](http://www.instytutobywatelski.pl))

Instytut Sobieskiego ([www.sobieski.org.pl](http://www.sobieski.org.pl))

Instytut Studiów Strategicznych ([www.iss.krakow.pl](http://www.iss.krakow.pl))

Instytut Spraw Publicznych ([www.isp.org.pl](http://www.isp.org.pl))

Krytyka Polityczna ([www.krytykapolityczna.pl](http://www.krytykapolityczna.pl))

Ośrodek Myśli Politycznej (<http://www.omp.org.pl>)

Ośrodek Studiów Wschodnich ([www.osw.waw.pl](http://www.osw.waw.pl))

Polski Instytut Studiów Międzynarodowych ([www.pism.pl](http://www.pism.pl))

Think Tank (<http://www.mttp.pl>)

Wiele z tych organizacji przygotowuje raporty i prowadzi badania. Bardzo często bezpłatnie je udostępnia czy wręcz aktywnie rozsyła do zainteresowanych. Te dane stanowią źródło opinii i wzbogacają debatę publiczną.

Różny jest poziom autorytetu tych instytucji i ich niezależności. Jak pisaliśmy wyżej z definicji mogą być one wprost zapleczem dla działań struktur politycznych, co absolutnie nie jest rzeczą naganną. Ważne jednak, żeby fakt ten był znany, a nie kamuflowany.

Część z nich tworzonych jest w formie tzw. QNGO (*quasi non governmental organization*), a zatem przez struktury publiczne. Polskim przykładem może być wymieniony wyżej znany i bardzo ceniony Ośrodek Studiów Wschodnich.

W działaniach rzeczniczych instytucje typu think tank mogą być ważnym sojusznikiem, ale trzeba je dobrze dobrać, rozumiejąc konsekwencje współpracy, pamiętając, że często w procesie rzecznictwa ważne jest nie tylko, jakie fakty są przywoływane, ale także osoba tego, kto je formułuje.

W niektórych przypadkach wskazane jest też sięgnięcie do opracowań naukowych. Poszukiwać można ich przez portal Google Scholar (<http://scholar.google.pl>), ale niestety nie zawsze są one darmowe. W przypadku artykułów można uzyskać dostęp do baz pełnotekstowych przez ośrodki akademickie – z kolei fragmenty zagranicznych książek naukowych można „podejrzeć” w księgarni internetowej Amazon.com lub w serwisie Google Books. Czasem po przejrzeniu takiej próbki może się okazać, że warto zainwestować w zakup tego typu publikacji lub przynajmniej czasowy dostęp do jej wersji elektronicznej.

### **Plusy i minusy wykorzystania danych tego typu**

- + Raporty stanowią gotową syntezę, nie wymagającą dodatkowej obróbki. Z reguły łatwo wyciągnąć z nich najbardziej interesujące nas informacje, oczywiście przywołując odpowiednie źródło.
- + Korzystanie z danych zawartych w raportach nie wymaga ani dużych środków, ani dużego nakładu pracy.
- + Przywołanie istniejących opracowań pokazuje, że jesteśmy zorientowani w danym temacie.
- W części przypadków dane zawarte w istniejących opracowaniach są już znane i nie wprowadzają nic nowego do debaty nad daną kwestią.
- Powoływanie się wyłącznie na cudze raporty nie buduje wizerunku naszej organizacji jako specjalizującej się w określonej dziedzinie.



- Korzystanie z danych i opracowań zapożyczonych od innych instytucji w pewnych sytuacjach wzmacnia naszą argumentację (w przypadku źródeł o niekwestionowanym autorytecie), ale bywa też, że osłabia lub komplikuje naszą pozycję w stosunku do pewnych środowisk – szczególnie, gdy źródło ma wyraźne zabarwienie ideologiczne.

## 4.2. Dokumenty strategiczne

### Krótką charakterystyka

Dokumenty strategiczne, a więc różnego rodzaju długofalowe strategie opisujące cele rozwojowe, jak konkretne działania, które będą w związku z nimi podejmowane, mogą być ważnym źródłem danych w działalności rzeczniczej. Strategie mogą być tworzone na poziomie ogólnokrajowym, jak i regionalnym, powiatowym i lokalnym. Mogą mieć charakter przekrojowy, wielosektorowy (np. strategie rozwoju gminy, ale także bardziej tematyczny (np. strategia nastawiona na poprawę sytuacji pewnej konkretnej grupy).

Pozwalają na określenie, jak na daną kwestię patrzy administracja, z jakich danych sama korzysta i co proponuje. Większość dokumentów strategicznych zawiera również część diagnostyczną, którą można potraktować jako osobne źródło, analogiczne do raportu.

Oczywiście dokumenty strategiczne (ich zawartość) same w sobie mogą być też przedmiotem działań rzeczniczych. Coraz częściej bowiem w ich tworzenie angażowane są różnego rodzaju podmioty, w tym organizacje pozarządowe i sami mieszkańcy. Powstawaniu dokumentów towarzyszą konsultacje publiczne, w których powinny brać i biorą udział także organizacje rzecznicze.

### Kiedy z nich korzystać?

Możemy powoływać się na dokumenty strategiczne, kiedy chcemy podkreślić systemową wagę spraw, którymi się zajmujemy. Co więcej, tego typu dokumenty mogą służyć do mobilizowania władz do działania – znajdują się w nich bowiem mniej lub bardziej konkretne deklaracje, które rządzący realizują bądź nie. Można też odwoływać się do strategii wyższego rzędu, np. Strategii Lizbońskiej, oceniając działania władz krajowych.

Możemy korzystać z dokumentów strategicznych również w sytuacji, w której chcemy dowiedzieć się, na jakich źródłach danych opiera się administracja. Wreszcie, dość często wskazywanie na ich zapisy pozwala unaocznnić niekonsekwencje czy zaniechania w działaniach władz. Jest to dość bolesny, a przez to często skuteczny, argument w debacie publicznej.

### **Gdzie je znaleźć?**

Dokumenty strategiczne są ogólnodostępne i z reguły znajdują się na stronach internetowych (bądź w Biuletynach Informacji Publicznej) instytucji odpowiedzialnych za ich przygotowanie. Na poziomie krajowym to na ogół poszczególne ministerstwa, na poziomie samorządowym – odpowiednie urzędy.

### **Plusy i minusy wykorzystania danych tego typu**

- + Dokumenty strategiczne są najczęściej bardzo łatwo dostępne (a w każdym razie powinny być).
- + Tego typu dokumenty pozwalają rozliczać władze z realizacji zobowiązań.
- + Administracja wykorzystuje często własne dane, których nie udostępnia w innych miejscach – jest to więc dobry sposób, by do nich dotrzeć. Dane te są też siłą rzeczy trudne do podważania przez samą administrację.
- Część diagnostyczna w dokumentach strategicznych jest nierzadko zaniedbywana i ma niską jakość. Jest też czasami słabo powiązana z priorytetami strategii.
- Dokumenty strategiczne mają z reguły dość duży poziom ogólności.

## **4.3. Statystyka publiczna**

### **Krótką charakterystyka**

Statystyka publiczna jest podstawowym źródłem danych wykorzystywanych w debatach toczonych przy tworzeniu polityk publicznych. Znajomość „tajników” pracy Głównego Urzędu Statystycznego i informacji przez niego gromadzonych stanowi w istocie niemal odrębną dziedzinę wiedzy, dlatego też poświęcamy jej znaczną część kolejnego rozdziału podręcznika [patrz podrozdział 5.3.1 „Główny Urząd Statystyczny”, str. 44]. W tym miejscu skupiamy się tylko na podstawowych cechach dotyczących statystyki publicznej jako źródła danych.

### **Kiedy z nich korzystać?**

Statystyka publiczna pozwala na poznanie „oficjalnego” obrazu sytuacji w danej dziedzinie. To, jakie problemy są rejestrowane i za pomocą jakich wskaźników są mierzone, w znacznym stopniu determinuje znaczenie zgromadzonych danych. Można wykorzystać tę sytuację na dwa sposoby. Z jednej strony, statystyka publiczna dostarcza danych, które raczej nie będą (lub nie powinny być) kwestionowane przez administrację. Z drugiej strony, można podkreślać właśnie to, że niektóre zjawiska mierzone są w niewłaściwy sposób i dążyć do zmiany rozmaitych definicji wykorzystywanych przez urzędy statystyczne.

## Gdzie je znaleźć?

Program statystyki publicznej realizowany jest w Polsce przez Główny Urząd Statystyczny oraz przez urzędy statystyczne w województwach, które nie tylko gromadzą dane regionalne, ale również specjalizują się w określonych dziedzinach. Nie należy zapominać o statystykach z poziomu ponadnarodowego, za które w Unii Europejskiej odpowiedzialny jest Eurostat. Zarówno GUS, jak i Eurostat udostępniają swoje zbiory w postaci dostępnych w Internecie baz danych oraz publikacji – niektóre z nich są darmowe, za inne trzeba zapłacić.

## Plusy i minusy wykorzystania danych tego typu

- + Większość danych jest ogólnodostępna i bezpłatna.
- + Gromadzone dane dotyczą różnych dziedzin, różnych obszarów terytorialnych i różnych okresów, co pozwala na rozmaite porównania.
- + Dane te są oficjalne, co zwiększa ich wagę argumentacyjną (zwłaszcza w relacjach z administracją publiczną).
- + Badania GUS są robione regularnie i na dużych próbach.
- Są to najczęściej wykorzystywane dane, dlatego rzadko wnoszą coś całkowicie nowego do dyskusji.
- Często wykorzystywane kategorie lub sposoby zbierania danych mogą wprowadzać w błąd lub utrudniać porównanie z innymi danymi.
- Duża ilość danych zgromadzonych przez urzędy statystyczne jest trudno dostępna ze względu na przepisy dotyczące tajemnicy statystycznej lub wysoką cenę udostępnienia.

## 4.4. Pozostałe dane o charakterze rejestrowym

### Krótką charakterystyka

Nie tylko Urzędy Statystyczne gromadzą dane o charakterze rejestrowym. Wiele instytucji publicznych przy okazji swoich działań zbiera informacje, o których często mało kto wie, a które mogą być niezwykle przydatne dla organizacji rzeczniczych. O tego typu danych piszemy więcej w kolejnym rozdziale podręcznika **[patrz podrozdział 5.1 „Dane publiczne”, str. 38]**.

### Kiedy z nich korzystać?

Dane administracyjne mogą odgrywać podobną rolę, jak statystyka publiczna. Poza tym mają jednak jeszcze jedną dodatkową cechę. Często nie są powszechnie dostępne, ich ujawnienie może więc rzucać nowe światło na zagadnienie, które jest przedmiotem dyskusji.

## Gdzie je znaleźć?

Wiele organów administracji rządowej i samorządowej jest często w posiadaniu danych, z których nie zawsze wystarczająco korzystają. Najprostszym sposobem ich zdobycia – o ile nie są dostępne na stronach internetowych urzędów – jest tryb dostępu do informacji publicznej.

Można także skorzystać z repozytoriów stworzonych przez organizację, które zajmują się pozyskiwaniem tego typu danych i ich rozpowszechnianiem. Stowarzyszenie Klon/Jawor na podstawie materiałów uzyskiwanych z różnych urzędów stworzyło serwisy [mojapolis.pl](http://mojapolis.pl) oraz [naszakasa.org.pl](http://naszakasa.org.pl) – oba opisujemy w dalszej części podręcznika [patrz podrozdziały 5.3.5 i 5.3.6, str. 54 i 60]. Ponadto Fundacja ePaństwo prowadzi serwis [www.sejmometr.pl](http://www.sejmometr.pl).

Czasami istotna jest sama wiedza o tym, że pewne dane istnieją, nawet jeśli nie mamy do nich dostępu. Pomocny może być w tym wykaz prawie 600 systemów informacyjnych administracji publicznej, który można znaleźć na stronie GUS<sup>2</sup>.

Dane rejestrowe nie są tylko i wyłącznie domeną administracji publicznej. Czasem ich gromadzeniem zajmują się również instytucje prywatne, w tym wiele organizacji pozarządowych. Przykładem może być tu baza osób zaginionych Fundacji Itaka. W niektórych przypadkach tego typu bazy danych organizacji są jedynymi dostępnymi materiałami.

## Plusy i minusy wykorzystania tego typu danych

- + Wykorzystanie danych, które dotychczas były niedostępne, pozwala na świeże spojrzenie i wprowadzenie do debaty nowych argumentów.
- + W niektórych mało dyskutowanych kwestiach są to jedyne istniejące dane.
- + Tego typu dane dają często duże możliwości dezagregacji i pozwalają na analizę na dużym poziomie szczegółowości.
- Dane przechowywane przez administrację bywają niskiej jakości – są często nieuporządkowane, słabo opisane, mogą zawierać dużo błędów albo przekłamań np. gdy tworzone są jako mechanizm sprawozdawczy o własnych działaniach i ich skuteczności.
- Z tego powodu doprowadzenie danych do stanu, w którym mogą być użyteczne, wymaga dużo pracy i wiedzy dotyczącej tego, jak różnego rodzaju dane są konstruowane.

---

2 [http://www.stat.gov.pl/bip/2815\\_PLK\\_HTML.htm](http://www.stat.gov.pl/bip/2815_PLK_HTML.htm)

## 4.5. Akty prawne i projekty aktów prawnych

### Krótką charakterystyka

Kiedy chcemy wpływać na tworzenie polityk publicznych i kształtować debatę o danym problemie, jednymi z kluczowych dokumentów są obowiązujące i planowane przepisy prawne. Dzieje się tak, nawet jeśli uznajemy, że zmiana społeczna niekoniecznie dokonuje się poprzez zmiany w prawie. Przepisy mówią o punkcie wyjścia działań i o polu możliwości – pozwalają też spojrzeć na problem z punktu widzenia, który jest bliższy przedstawicielom administracji. Dodatkowym ważnym źródłem informacji mogą być uzasadnienia danych inicjatyw legislacyjnych oraz tzw. oceny skutków regulacji.

Nie bez znaczenia jest też fakt, że kształt owych przepisów (jego zmiana) bywa zasadniczym celem działań rzeczniczych. Z tego punktu widzenia warto odnotować fakt, że coraz częściej prawo do uczestnictwa w ich konsultowaniu przez organizacje nie jest kwestionowane, a wręcz staje się regułą. Za ważne można uznać przyjęcie przez rząd w 2012 r. w ramach programu Lepsze Regulacje tzw. siedem zasad konsultacji (<https://mac.gov.pl/konsultacje-spoeczne>), które obecnie wdrażane są stopniowo w poszczególnych resortach (pierwszym z nich było Ministerstwo Administracji i Cyfryzacji).

### Kiedy z nich korzystać?

Na przepisach warto skupić się w szczególności wtedy, gdy można mieć podejrzenia, że zmiana prawa daje szansę na realną zmianę w temacie, w którym prowadzimy działania rzecznicze.

### Gdzie je znaleźć?

Uchwalone akty prawne z poziomu centralnego dostępne są w Internetowym Systemie Aktów Prawnych (<http://isap.sejm.gov.pl>), który gromadzi zawartość Dziennika Ustaw i Monitora Polskiego.

Projekty rozproszone są po różnych miejscach, zależnie od etapu procesu legislacyjnego. Część znaleźć można na stronach Sejmu, część w Biuletynie Informacji Publicznej Rady Ministrów. Niektóre projekty pojawiają się na internetowych platformach konsultacyjnych, takich jak MamZdanie.org.pl. Ważnym adresem jest także platforma rządowa [www.konsultacje.gov.pl](http://www.konsultacje.gov.pl), która docelowo ma być jednym wspólnym miejscem, w którym będą upubliczniane i poddawane konsultacjom publicznym wszystkie rządowe projekty zmian legislacyjnych. Aktualnie z platformy tej pilotażowo korzysta Ministerstwo Gospodarki.

Istnieje również przywoływany już serwis [www.sejmometr.pl](http://www.sejmometr.pl), gromadzący m.in. dane o procesie legislacyjnym prowadzonym w ramach prac parlamentu.

Na poziomie samorządowym wszystkie prace nad prawem miejscowym powinny być odzwierciedlone w Biuletynach Informacji Publicznej, niestety bardzo często są one niekompletne.

### **Plusy i minusy wykorzystania danych tego typu**

- + Argumentacja odwołująca się do konkretnych rozwiązań prawnych jest przekonująca dla prawodawców i/lub administracji, która wręcz nie może ignorować argumentów i rozwiązań zakotwiczonych w przepisach prawnych.
- + Proces tworzenia prawa coraz częściej formalnie podlega procesom konsultacji publicznych. To dobry moment na wyrażenie swojej opinii i domagania się reakcji na nią.
- Dla wielu odbiorców debata na poziomie rozwiązań prawnych może być nieatrakcyjna, zbyt techniczna i w ich mniemaniu odległa od realnych problemów, z którymi się borykają.
- Często zmiany w prawie okazują się tylko pozornie skuteczne, a prawo pozostaje martwe.
- Dokładne zapoznanie się z zawiłościami przepisów jest czasochłonne i wymaga dużej wiedzy prawniczej, łatwo też o pomyłkę.

## **4.6. Przebieg procesu legislacyjnego**

### **Krótką charakterystyka**

Istotne jest nie tylko, jakie prawo powstaje, ale również, jak ono powstaje. Poza nielicznymi nagłaśnianymi przez media skrajnymi przypadkami przeciętni obywatele niewiele wiedzą o szczegółowym przebiegu stanowienia prawa. Ukazywanie w konkretnej sytuacji, kto podjął jaką decyzję i w jakich okolicznościach określony zapis znalazł się w nowej ustawie, może mieć duże znaczenie w debacie o danym problemie.

### **Kiedy z nich korzystać?**

Analizowanie procesu stanowienia prawa może służyć kontroli, czy przebiega on zgodnie z określonymi standardami. Możemy ustalić, kto ma wpływ na powstanie danego rozwiązania, albo czy sprzyja ono jakimś interesom lub je narusza.

### **Gdzie je znaleźć?**

Źródła danych o stanowieniu prawa są podobne, jak w przypadku samego prawa. Ważne są zwłaszcza dokumenty opisujące sam proces, np. protokoły z posiedzeń organów, takich jak komisje sejmowe czy rada gminy.

Szczególnie ważnym źródłem informacji o procesie legislacyjnym prowadzonym w ramach struktur rządowych jest serwis Rządowego Centrum Legislacji (<http://legislacja.rcl.gov.pl>). Można się tam zapoznać z aktualnymi i planowanymi aktami prawa (nie tylko z ustawami, ale także z rozporządzeniami) w fazie, w której się aktualnie znajdują oraz z towarzyszącą im dokumentacją. Na etapie prac parlamentarnych proces ten opisywany jest przez kanał sejmowy (<http://www.sejm.gov.pl/>). W jego ramach funkcjonuje też serwis video iTVSejm, za pomocą którego można oglądać bieżące i archiwalne posiedzenia plenarne, a także niektóre posiedzenia Komisji. Te ostatnie mogą mieć o wiele większe znaczenie z punktu widzenia kształtu legislacji.

W tym kontekście warto też wymienić prowadzone przez same organizacje pozarządowe serwisy: [MamPrawoWiedziec.pl](http://MamPrawoWiedziec.pl), który pomaga śledzić proces uchwalania prawa w parlamencie, poprzez ukazywanie w przejrzysty sposób, jak głosują poszczególni posłowie i senatorowie w określonych sytuacjach, oraz spełniający podobną funkcję wymieniony już wcześniej serwis [www.sejmometr.pl](http://www.sejmometr.pl).

### **Plusy i minusy wykorzystania danych tego typu**

- + „Sprawczość” dokumentów legislacyjnych, a co za tym idzie ich szczególna waga z punktu widzenia działań rzeczniczych.
- + Powszechna dostępność i na ogół dobra strukturyzacja, zarówno dokumentów, jak i powiązanych z nim danych i opracowań.
- + Możliwość egzekwowania dokumentów (i powiązanych z nimi opracowań) na gruncie oczywistego postulatu jawności procesu legislacyjnego.
- + Możliwość powiązania kształtu dokumentów z ich autorami oraz uczestnikami procesu decyzyjnego (identyfikacji indywidualnych głosowań/wypowiedzi, a także uczestników działań konsultacyjnych), a co za tym idzie lepsze zrozumienie mapy interesariuszy i możliwość precyzyjniejszego adresowania działań rzeczniczych.
- Częste trudności w nawigacji, mnogość dokumentów, brak tagowania, złe formaty (np. nieprzeszukiwalny PDF), hermetyczny język, brak ujednolicenia tekstów etc.
- Nierzadkie przypadki zaniechania konsultacji albo ignorowania ich wyników.
- Szybkie i zaskakujące zmiany kształtu dokumentów niemożliwe do wychycenia bez odpowiednich zasobów na ich monitorowanie i na nie reagowanie.

## 4.7. Informacje o kluczowych aktorach

### Krótką charakterystyka

W wielu dziedzinach, zarówno na poziomie krajowym, jak i lokalnym, funkcjonują osoby lub grupy osób kluczowe z punktu widzenia interesujących nas działań [patrz podrozdział 3.3., „Analiza interesariuszy”, str. 20]. Tacy aktorzy mogą być ważni ze względu na możliwość podjęcia pewnej decyzji, nagłośnienia jakiejś sprawy czy też wywarcia nacisku na kogoś innego – lub wręcz przeciwnie, z powodu promowania stanowiska odmiennego od naszego i bycia potencjalnymi oponentami.

### Kiedy z nich korzystać?

Pozyskiwanie informacji o osobach, grupach czy instytucjach może polegać na poszukiwaniu powiązań między różnymi aktorami zaangażowanymi w proces stanowienia prawa. Dobra znajomość poglądów i działalności określonych postaci umożliwi również wykazywanie im sprzeczności lub niekonsekwencji. Lub wręcz przeciwnie, można też tę wiedzę wykorzystać do zainteresowania kogoś tematem, który jest mu bliski.

### Gdzie je znaleźć?

Większość danych tego typu wymaga poszukiwań na własną rękę, poprzez przeczesywanie dokumentów i stron internetowych, które mogą dotyczyć danych aktorów. Należy pamiętać, że ta metoda naraża nas na ryzyko zdobycia informacji mało wiarygodnych czy niesprawdzonych.

Jest jednak kilka miejsc, gdzie można takich informacji szukać w sposób bardziej systematyczny. W Krajowym Rejestrze Sądowym można odnajdywać powiązania z przedsiębiorstwami i organizacjami, a w Państwowej Komisji Wyborczej można znaleźć informacje dotyczące aktywności wyborczej. Natomiast Serwis [MamPrawoWiedziec.pl](http://MamPrawoWiedziec.pl) zbiera w jednym miejscu wiele informacji o osobach piastujących wybieralne funkcje publiczne, dotyczących zarówno działań, jak i poglądów – jednak na razie obejmuje głównie parlamentarzystów. Poza dostępnymi publicznie w różnej formie danymi na temat np. sposobu głosowania czy oświadczeń majątkowych w serwisie gromadzone są odpowiedzi na specjalnie przygotowany kwestionariusz, który zbiera dane na temat preferencji poszczególnych kandydatów z okresu zanim zostali wybrani (chodzi o to, aby pozwolić wyborcom dokonywać selekcji w bardziej z ich punktu widzenia racjonalny sposób). Kwestionariusz tworzony jest we współpracy z kluczowymi organizacjami działającymi w poszczególnych obszarach tematycznych. W oparciu o jego wyniki można również poszukiwać sojuszników lub lokalizować oponentów wśród parlamentarzystów lub kandydatów.



### **Plusy i minusy wykorzystania danych tego typu**

- + Dane tego typu pozwalają opowiedzieć historię, która jest atrakcyjna dla opinii publicznej.
- + W wielu działaniach kluczem do sukcesu jest znalezienie decydentów, którzy będą się identyfikowali z naszymi propozycjami (wręcz potraktują je jak własne).
- Wykorzystywanie argumentów odnoszących się do konkretnych osób bywa przeciwnie skuteczne w działaniach systemowych.
- Wytykanie komuś działań lub poglądów (ustalanych nie zawsze w oparciu o wiarygodne źródła) może łatwo prowadzić do konfliktu. Spór personalny nabiera wtedy autonomicznego charakteru, a argumenty merytoryczne tracą na znaczeniu.

## **4.8. Monitoring mediów**

### **Krótką charakterystyka**

Zawartości mediów nie można określić jako „twardych” danych – jednak ona również może być dla nas ważna. Przedmiotem analizy jest w tym przypadku przede wszystkim sposób wypowiadania się o danych problemach. Przekazy medialne (najczęściej prasowe) są traktowane jako odzwierciedlenie stanu debaty publicznej.

### **Kiedy z nich korzystać?**

Dane tego typu z reguły nie mogą być argumentem w debacie, ale pozwalają przedstawić tło danego problemu i ustalić stanowiska różnych stron. Pomagają również opisać pozycję mediów jako jednego z aktorów.

### **Gdzie je znaleźć?**

Monitoring prasy w podstawowym zakresie jest bardzo prosty, ponieważ większość tytułów posiada archiwa internetowe, część z nich jest nawet bezpłatna. Większym problemem jest analiza materiałów telewizyjnych lub radiowych, nie tylko ze względu na dostęp, ale również przez to, że trzeba uwzględnić więcej czynników (różne elementy przekazu dźwiękowego i obrazu).

Istnieją też narzędzia – w tym darmowe, jak np. Google Alerts – pozwalające monitorować zasoby internetowe pod kątem pojawiania się treści dotyczących interesującego nas zagadnienia.

### **Plusy i minusy wykorzystania danych tego typu**

- + Monitoring prasy jest dobrym źródłem informacji o trendach w debacie na dany temat.

- + Dane tego typu pozwalają poznać stanowiska różnych stron.
- + Pozwalają też wpływać na język debaty poprzez wskazywanie na niewłaściwe treści.
- Dane te są często powierzchowne.
- Nie są one twardym argumentem, którym można łatwo kogoś przekonać.

## 4.9. Świadcstwa

### Krótką charakterystyka

Czasami zamiast danych omawiających ogólnie dany problem, bardziej znaczące są jednostkowe historie obrazujące go. Mogą to być opisy doświadczeń konkretnych ludzi czy poszczególnych wydarzeń, często wykorzystuje się też materiały wizualne, które oddziałują szczególnie mocno. Nie dowodzą one skali problemu, ale pokazują, na czym on polega. Takie dane są bardzo chętnie wykorzystywane przez organizacje.

### Kiedy z nich korzystać?

Twarde dane są istotne w dyskusji między specjalistami zajmującymi się daną dziedziną. Jednak kiedy kierujemy nasz przekaz do opinii publicznej, wiele osób jest bardziej zainteresowanych konkretnymi przykładami, które pokażą, o co nam chodzi. Świadcstwa przemawiają bardziej do emocji niż stanowią racjonalny argument. W związku z tym mogą być elementem perswazji, kiedy chcemy mobilizować do działania. Te dane są ponadto często wykorzystywane, gdy brakuje nam już innych.

### Gdzie je znaleźć?

Większość organizacji, zwłaszcza świadczących usługi, korzysta w tym wypadku z własnych doświadczeń, gromadząc historie ludzi, z którymi pracuje. Świadcstwa mogą być również produktem pogłębionych badań jakościowych, nastawionych na poszukiwanie przypadków szczególnie ważnych z punktu widzenia problemu, którym się zajmujemy.

### Plusy i minusy wykorzystania danych tego typu

- + Świadcstwa działają na emocje, są przekonujące.
- + O ile są zgodne z danymi liczbowymi, świadcstwa potrafią je bardzo silnie wzmocnić i uwiarygodnić. To samo działa w drugą stronę – trudno podważyć wnioski płynące z pojedynczego przypadku, jeśli liczby go potwierdzają.
- Danym tego typu nie można uogólniać i wnioskować na ich podstawie o tym, jak wygląda całościowy problem.

- Świadczenia bardzo łatwo podważyć, choćby przywołując przypadek dowodzący czegoś odwrotnego niż ten, którym się posłużyliśmy.
- Z wielką delikatnością trzeba podchodzić do opowiadania indywidualnych historii. Zadbaj o to, aby nie urazić prywatności i godności osób, których dotyczą. Warto też pamiętać, że za naruszenie dóbr osobistych (w tym prywatności) grożą sankcje prawne. Jeśli chcemy używać czyjegoś wizerunku lub historii, nie powinniśmy tego czynić bez zgody samych zainteresowanych.

## 4.10. Wyniki badań społecznych

### Krótką charakterystyka

Istnieje wiele danych tego rodzaju. Ich wytwarzaniem zajmują się różnego rodzaju instytucje, zarówno publiczne (przede wszystkim przywoływany już Główny Urząd Statystyczny oraz Centrum Badania Opinii Społecznej, ale np. też Państwowy Fundusz Rehabilitacji Osób Niepełnosprawnych), czynią to czasem także instytucje regionalne np. Regionalne Obserwatoria Polityk Społecznych czy tzw. Obserwatoria Rynku Pracy. Badania takie prowadzi także nierzadko samorządy (np. warszawskie badania na próbie ok. 9 tys. mieszkańców z 2011 r.). W związku z tym, że ich wykonanie opłacane jest z podatków, dostęp do nich powinien być łatwiejszy niż do innych źródeł, jak te prywatne. Więcej na temat tego rodzaju danych piszemy w dalszej części podręcznika [\[patrz podrozdział 5.3.7 „Dostępne wyniki badań społecznych”; str. 66\]](#).

Powstaje także wiele badań na poziomie międzynarodowym (np. Światowy Sondaż Wartości, Europejski Sondaż Społeczny).

Wyniki badań społecznych są użyteczne na dwa sposoby. Po pierwsze, możemy powoływać się na gotowe raporty. Przedstawione w nich wyniki mogą być argumentem dla organizacji rzeczniczych. Po drugie, niektóre instytucje tworzące opracowania upubliczniają również dane surowe (np. wyniki Diagnostyki Społecznej). Daje to możliwość dokonania pogłębionej analizy materiału. Dotyczy to jednak niewielu badań.

### Kiedy z nich korzystać?

Tego rodzaju dane pozwalają zgromadzić informacje o poglądach, postawach i zachowaniach pewnej populacji. Są z reguły łatwo dostępne, zatem są użyteczne podczas poszukiwania informacji o interesującym nas zagadnieniu.

### **Skąd można je zdobyć?**

Wiele tego rodzaju danych publikowanych jest na stronach internetowych jednostek, które je tworzyły. W przypadku badań prowadzonych przez administrację publiczną, których wyniki nie są osiągalne, niezbędna jest próba pozyskania ich poprzez dostęp do informacji publicznej.

### **Plusy i minusy wykorzystania danych tego typu**

- + Korzystanie z istniejących już wyników jest tanie, pozwala zaoszczędzić duże środki. Dotyczy to jednak tylko części badań – w większości tych, które są prowadzone ze środków publicznych.
- + O ile badania społeczne zostały poprawnie przeprowadzone, możemy większość danych generalizować na całą badaną populację.
- Podstawową wadą tego typu danych jest łatwość ich podważania, zwłaszcza w przypadku badań sondażowych. Nie mają dużego ciężaru dowodowego, często nie są uznawane za wiarygodne.
- Aby samodzielnie analizować surowe dane (a często tylko to może przynieść interesujące nas informacje), potrzebna jest duża wiedza i kompetencje statystyczne. Co więcej, póki co te dane udostępniane są sporadycznie (wyjątkiem jest np. Diagnoza Społeczna).

## **4.11. Badania własne i zleczone**

### **Krótką charakterystyka**

Najważniejszą cechą badań własnych jest z pewnością to, że organizacja ma niemal nieograniczone (limitem jest wyłącznie budżet) możliwości formułowania pytań, na które chce poznać odpowiedź. Realizacja badań tworzy również określony wizerunek organizacji, która zna się na danej dziedzinie.

### **Kiedy z nich korzystać?**

Badania własne i zleczone realizuje się zazwyczaj wtedy, gdy po etapie analizy danych zastanych okazuje się, że brakuje informacji na interesujący nas temat. Często też dochodzimy do wniosku, że dostępne dane są nieaktualne lub mało wiarygodne i chcemy skonfrontować je z własnymi badaniami.

### **Skąd można je zdobyć?**

Są dwie główne możliwości pozyskania tego rodzaju danych. Z jednej strony można je zlecić zewnętrznej jednostce badawczej, która specjalizuje się w realizacji badań z określonej dzie-

dziny. Jest to jednak rozwiązanie bardzo kosztowne. Co więcej, w przypadku nieporozumienia z instytucją badawczą, można otrzymać wyniki nieodpowiadające naszym oczekiwaniom. Drugą możliwością jest samodzielne realizowanie badania. Ta opcja wymaga od organizacji również sporych nakładów finansowych. To jednak nie wystarczy – organizacja powinna mieć ponadto dobrze przemyślany pomysł na badanie i kompetentne osoby w zespole lub wśród współpracowników, które będą w stanie takie badanie przeprowadzić.

### **Plusy i minusy wykorzystania danych tego typu**

- + Własne badania pozwalają na wytworzenie nowej wiedzy o danej dziedzinie, uzupełnienie istniejących luk.
- + Dostarczanie nowej wiedzy pozwala tworzyć prestiż organizacji będącej ekspertem w danej dziedzinie.
- Własne badania są z reguły kosztowne, zwłaszcza jeśli ich realizację chcemy zlecić profesjonalnej instytucji.
- Wymagają również dużej wiedzy. Nawet jeśli chcemy zlecić badanie na zewnątrz, potrzebujemy podstawowego zasobu wiadomości o badaniach, żeby precyzyjnie umówić się z instytucją badawczą. Dodatkowym etapem wymagającym kompetencji jest analiza zebranego materiału.

## 5. Zasoby danych dla działalności rzeczniczej w Polsce

W poprzednim rozdziale podkreślaliśmy, że dane można rozumieć szerzej niż statystyki – również jako dokumenty, historie ludzi, przekazy prasowe. Jednak w dalszej części podręcznika skupimy się mimo wszystko na danych ilościowych – one bowiem wymagają pewnych konkretnych umiejętności. Uważane powszechnie za „obiektywne”, są najbardziej narażone na ryzyko manipulacji. Jednocześnie powinny stanowić ważny argument w dyskusji z decydentami – a niekoniecznie zawsze tak jest.

Osoby, chcące dowiedzieć się więcej o skutecznym wykorzystaniu miękkich materiałów w działaniach rzeczniczych, zachęcamy do zapoznania się np. z publikacjami Tactical Technology Collective ([www.tacticaltech.org](http://www.tacticaltech.org)). TTC to międzynarodowa organizacja wspierająca wykorzystanie nowych technologii w rzecznictwie. W publikacjach z cyklu „10 Tactics” (<http://informationactivism.org/en>) przywołuje się wiele organizacji z całego świata posiłkujących się w swoich działaniach różnego rodzaju informacjami czy też danymi (zarówno ilościowymi, jak i innymi). Niestety wszystkie te źródła są wyłącznie w języku angielskim.

### 5.1. Dane publiczne

Szczególnie ważną kategorią, której chcielibyśmy się bliżej przyjrzeć, są dane publiczne. Jest to grupa szeroka, obejmująca bardzo różnorodne typy informacji. Dane publiczne będziemy tu rozumieć tak, jak zdefiniowało je Stowarzyszenie Klon/Jawor w niedawno opublikowanym raporcie pt. „Dostęp do danych publicznych w Polsce”<sup>3</sup>, czyli jako dane pozostające pod kontrolą administracji publicznej. Z jednej strony chodzi więc o materiały będące w posiadaniu GUS, a z drugiej – te gromadzone przez inne urzędy, na poziomie centralnym oraz na niższych poziomach.

Głównym Urzędem Statystycznym, jako instytucją z założenia zajmującą się gromadzeniem, przetwarzaniem i udostępnianiem informacji, zajmiemy się szczegółowo w osobnym podrozdziale **[patrz podrozdział 5.3.1 „Główny Urząd Statystyczny”, str. 44]**. Warto jednak zatrzymać

---

3 [http://www.mojapolis.pl/c/document\\_library/get\\_file?uuid=5333cfe8-a7e2-4a84-9d3b-2a1dac8d3949&groupId=10122](http://www.mojapolis.pl/c/document_library/get_file?uuid=5333cfe8-a7e2-4a84-9d3b-2a1dac8d3949&groupId=10122)

się na chwilę przy kwestii pozyskiwania danych z innych organów administracji. Przy okazji typologii pisaliśmy [patrz podrozdział 4.4 „Pozostałe dane o charakterze rejestrowym”; str. 27], że archiwa instytucji państwowych i samorządowych mogą być źródłem wielu ciekawych informacji. Jednak droga do ich zdobycia może być dosyć długa.

Jak może wyglądać pozyskiwanie danych od organów administracji publicznej, pokazuje wspomniany wyżej raport Stowarzyszenia Klon/Jawor. Autorzy przedstawiają szczegółowo kolejne podejmowane przez siebie kroki, co stanowi pomocny przykład. Podstawowym trybem pozyskiwania materiałów jest w tym przypadku wnioski o dostęp do informacji publicznej. Ten sposób komunikacji z administracją rządzi się określonymi regułami, ma swoje ograniczenia i sprawdzone metody. Najlepszym źródłem wiedzy w tym temacie jest Pozarządowe Centrum Dostępu do Informacji Publicznej (<http://informacjapubliczna.org.pl>), prowadzone przez Sieć Obywatelską – Watchdog Polska (dawniej Stowarzyszenie Liderów Lokalnych Grup Obywatelskich).

Poniżej przedstawimy kilka podstawowych informacji na ten temat.

---

### **Dostęp do informacji publicznej (IP)**

- Prawo do uzyskiwania IP jest zapisane w Konstytucji RP (art. 61).
  - Prawo to obejmuje dostęp do dokumentów (poprzez Biuletyn Informacji Publicznej lub wnioski) oraz wstęp na posiedzenia kolegialnych organów władzy publicznej pochodzących z wyborów powszechnych.
  - IP to „informacje wytworzone przez władze publiczne oraz osoby pełniące funkcje publiczne oraz inne podmioty, które wykonują funkcje publiczne lub gospodarują mieniem publicznym (komunalnym lub Skarbu Państwa), jak również informacje odnoszące się do wspomnianych władz, osób i innych podmiotów, niezależnie od tego, przez kogo zostały wytworzone” (wyrok NSA z dnia 30 października 2002 r., II SA 181/02).
  - Prawo dostępu do IP przysługuje każdemu obywatelowi, nie wolno żądać od niego podania, z jakiego powodu chce zdobyć daną informację.
  - Instytucja, do której wpłynął wniosek o udostępnienie IP, ma obowiązek udzielić odpowiedzi w ciągu 14 dni.
  - Trudniej jest uzyskać dostęp do IP, jeśli będzie to informacja przetworzona (czyli taka, która wymaga obliczeń, analiz lub zestawień statystycznych) – wtedy konieczne może być wykazanie interesu publicznego i/lub opłacenie pracy związanej z wygenerowaniem takiej informacji.
  - Prawo dostępu do IP podlega ograniczeniu wyłącznie ze względu na ochronę praw i wolności innych osób oraz bezpieczeństwo państwa.
-

Żeby jednak wykorzystać możliwość dostępu do informacji publicznej i spróbować „wyciągnąć” z upatrzonego przez nas urzędu interesujące dane, musimy wiedzieć, co możemy tam znaleźć i o co pytać (lub czego szukać w Biuletynie Informacji Publicznej). Z pomocą przychodzi tutaj strona internetowa Głównego Urzędu Statystycznego, a dokładniej opis Polskiej Statystyki Publicznej<sup>4</sup>. W tym roku po raz pierwszy umieszczono tam dwa dokumenty: „Wykaz systemów informacyjnych administracji publicznej 2013” to lista 587 rejestrów, wraz z ich pełnymi nazwami oraz instytucjami odpowiedzialnymi. Oprócz niego jest również kolejny dokument, „Systemy informacyjne administracji publicznej 2013”, w którym opisany jest każdy system, zajmując mniej więcej jedną stronę. Znajdziemy tam m.in. zapisy o tym, jaki jest cel prowadzenia danego systemu, jaki jest zakres objętych nim informacji, jakie są źródła danych oraz jakie są formy ich udostępniania.

Opisane systemy dotyczą najróżniejszych dziedzin, od Systemu Informacji Oświatowej czy Systemu Informacji o Szkolnictwie Wyższym po bazę informacji o zdarzeniach radiacyjnych Państwowej Agencji Atomistyki czy Bank Danych Drogowych Generalnej Dyrekcji Dróg Krajowych i Autostrad.

## 5.2. Polska statystyka publiczna

Innym pomocnym materiałem umożliwiającym zapoznanie się z potencjalnymi źródłami publicznych danych dla działalności jest Program Statystyki Publicznej<sup>5</sup>, corocznie uchwalany przez Radę Ministrów (co oznacza, że jest publikowany w Dzienniku Ustaw). Znajdziemy w nim szczegółowy opis 243 badań, które zostały lub zostaną przeprowadzone w 2013 roku, z podziałem na dziedziny tematyczne. Dowiemy się, jaki jest zakres podmiotowy i przedmiotowy badania, kto przekaże dane statystyczne, w jakiej formie i w jakim terminie oraz jak będą wyglądać wyniki badania i w jaki sposób zostaną udostępnione. Podobne programy publikowane są corocznie.

To jednak nie wszystko. Choć ponad 700 stron charakterystyki systemów informacyjnych państwa i prawie 500 stron Programu Statystyki Publicznej robi imponujące wrażenie, to tak naprawdę, kiedy interesuje nas jeden bardzo konkretny obszar, te opisy i tak często okazują się zbyt ogólne. Niekoniecznie dowiemy się z nich, jakie konkretnie informacje mamy w praktyce do dyspozycji (czy też moglibyśmy mieć, gdyby udało nam się do nich uzyskać dostęp). Całe szczęście dokumenty,

---

4 [http://www.stat.gov.pl/bip/2815\\_PLK\\_HTML.htm](http://www.stat.gov.pl/bip/2815_PLK_HTML.htm)

5 [http://www.stat.gov.pl/cps/rde/xbcr/bip/BIP\\_PBSSP2013\\_DzU2012\\_poz\\_1391.pdf](http://www.stat.gov.pl/cps/rde/xbcr/bip/BIP_PBSSP2013_DzU2012_poz_1391.pdf)



na podstawie których GUS gromadzi swoje dane, również są publicznie dostępne. Dzięki temu możemy zdobyć wiedzę o każdej informacji, którą GUS ma w swoim posiadaniu.

Badania GUS można podzielić na dwie główne kategorie. Pierwsza to obowiązkowe sprawozdania, które różnego rodzaju podmioty samodzielnie dostarczają do odpowiednich organów administracji publicznej. W tym wypadku teoretycznie powinniśmy uzyskiwać dane na temat wszystkich istniejących podmiotów danego typu. Badania te przeprowadzane są na podstawie formularzy sprawozdawczych, które możemy ściągnąć z odpowiedniej strony na portalu GUS<sup>6</sup>. I tak na przykład z formularza SOF-I (Sprawozdanie z działalności fundacji, stowarzyszeń i innych organizacji społecznych) możemy wyczytać, że GUS gromadzi informacje o przynależności organizacji do formalnych lub nieformalnych porozumień organizacji na poziomie krajowym lub międzynarodowym.

Druga kategoria badań GUS to badania ankietowe. Są one realizowane na dużych, reprezentatywnych próbach osób, gospodarstw domowych lub gospodarstw rolnych **[patrz podrozdział 7.1.11. „Próba”, str. 85]**. W tym wypadku z reguły informacje są zbierane przez ankierów. Kwestionariusze, które wypełniają, również są dostępne na stronie GUS<sup>7</sup>. Możemy się stamtąd dowiedzieć np., że w ramach „Ankiety o zużyciu paliw i energii w gospodarstwach domowych” ankierzy GUS zbierają informacje m.in. o tym, w jaki sposób ogrzewa się mieszkania albo jakie są klasy efektywności energetycznej wykorzystywanych urządzeń elektrycznych. Są to szczegółowe dane, które mogą okazać się bardzo przydatne dla organizacji zajmujących się kwestiami ekologii i energetyki.

Dokładne zapoznanie się z formularzami sprawozdawczymi oraz kwestionariuszami badań ankietowych daje nam obszerną wiedzę na temat danych, które ma do dyspozycji GUS. Jest to punkt wyjścia do próby ich wydobycia. Podstawowe dane są z reguły dostępne poprzez witrynę internetową urzędu lub jego regionalnych oddziałów – czy to w postaci raportów i roczników, czy zestawień statystycznych, czy wreszcie w postaci interaktywnych baz danych, takich jak baza Demografia czy Bank Danych Lokalnych.

Nie wszystkie interesujące nas dane są jednak upubliczniane. Te można próbować uzyskać zwracając się bezpośrednio do urzędu. Niestety, praktyka pokazuje, że często jest to trudne, ponieważ GUS podlega różnym ograniczeniom w udostępnianiu danych. Jednym z nich jest

---

6 <http://form.stat.gov.pl/formularze/2013/index.htm>

7 <http://form.stat.gov.pl/BadaniaAnkietowe/2013/harmonogram.htm>

tw. tajemnica statystyczna. Działalność urzędów statystycznych jest regulowana odrębną ustawą o statystyce publicznej (Dz.U. 1995 Nr 88 poz. 439), w której (w art. 10) znajduje się przepis mówiący, iż „Zbierane i gromadzone w badaniach statystycznych statystyki publicznej dane indywidualne i dane osobowe są poufne i podlegają szczególnej ochronie; dane te mogą być wykorzystywane wyłącznie do opracowań, zestawień i analiz statystycznych oraz do tworzenia przez służby statystyki publicznej operatu do badań statystycznych prowadzonych przez te służby; udostępnianie lub wykorzystywanie danych indywidualnych i danych osobowych dla innych niż podane celów jest zabronione (tajemnica statystyczna)”. Ten właśnie zapis sprawia, iż GUS oraz jego agendy nie są skłonne udostępniać podmiotom zewnętrznym danych jednostkowych (na przykład informacji pochodzących od pojedynczych podmiotów). Choć można zrozumieć przesłanki takich rozstrzygnięć, sprawiają one, że GUS nie może udostępniać również danych, które według wielu powinny być jawne – np. informacji o zatrudnieniu w jednostkach administracji publicznej. Bardzo niechętnie udostępnia też tzw. mikrodane – czyli zanonimizowane (pozbawione cech umożliwiających identyfikację poszczególnych podmiotów) analityczne zbiory danych ze swoich badań. Takie informacje najłatwiej pozyskać, nawiązując współpracę z jednostkami naukowymi i akademickimi, które mogą je uzyskać z GUS. Można też próbować zdobyć je za pośrednictwem administracji publicznej – np. współpracującego z nami resortu. Tego rodzaju współpraca ma także jeszcze jeden walor – administracja publiczna może uzyskiwać nieopublikowane dane z GUS nieodpłatnie, podczas gdy od innych podmiotów oczekuje się zwykle opłat, niekiedy bardzo wysokich.

Poza zanonimizowanymi danymi jednostkowymi, możemy być także zainteresowani niestandardowymi (a więc niepublikowanymi w ramach tzw. programu badań statystyki publicznej) zestawieniami czy tablicami bazującymi na danych zbieranych przez GUS. Również takie zestawienia są przygotowywane za opłatą – o ile nie mamy sojusznika uprawnionego do uzyskiwania ich nieodpłatnie<sup>8</sup>, musimy być przygotowani na znaczne koszty. Ze względu na omówione już przepisy o tajemnicy statystycznej, nie otrzymamy także z GUS zestawień, które umożliwiłyby jednoznacznie identyfikację określonego podmiotu – np. takich, w których w pojedynczej komórce znajdowałyby się dane o mniej niż 3 podmiotach lub udział pojedynczego podmiotu w zestawieniu byłby większy niż ¼ całości. Zawilości związane z korzystaniem z tego trybu omawia wspomniany już raport Stowarzyszenia Klon/Jawor, opisujący różne praktyki związane z korzystaniem z danych publicznych i informacji publicznej. Warto przy tym zapoznać się z opinią prawną na temat informacji przetworzonej, przygotowaną specjalnie na potrzeby tego raportu.

---

8 Więcej informacji na ten temat – por. [http://www.stat.gov.pl/katow/19\\_PLK\\_HTML.htm](http://www.stat.gov.pl/katow/19_PLK_HTML.htm)

W niektórych przypadkach istnieje możliwość ominięcia części trudności w dostępie do danych GUS. Instytucja ta jest bowiem częścią Europejskiego Systemu Statystycznego i ma obowiązek przekazywać określone dane do Urzędu Statystycznego Unii Europejskiej (Eurostatu). Statystyka publiczna funkcjonuje bowiem nie tylko na poziomie lokalnym czy krajowym, ale również międzynarodowym – i jest tam bardzo rozwinięta. Rozbudowanymi bazami danych dysponują instytucje europejskie – przede wszystkim Eurostat – oraz globalne. Koniecznie trzeba wspomnieć tu o Banku Światowym oraz Organizacji Współpracy Gospodarczej i Rozwoju (OECD). Gromadzą one wiele statystyk dotyczących rozwoju społecznego. Ważnym źródłem danych odnoszących się do określonych dziedzin są też agencje ONZ, takie jak UNDP czy UNESCO. Niektóre informacje dotyczące Polski łatwiej znaleźć w bazach tych organizacji niż w zasobach GUS.

Dane międzynarodowe są zapewne wykorzystywane stosunkowo rzadziej przez polskie organizacje rzecznicze. Jednak nawet dla organizacji zajmujących się kwestiami krajowymi lub lokalnymi mogą być czasami przydatne – przede wszystkim ze względu na możliwość opisanego za ich pomocą szerszego kontekstu procesów analizowanych na poziomie krajowym. Dają możliwość np. przyjrzenia się, dlaczego określonemu państwu udało się osiągnąć cel, do którego my również dążymy. W dyskusjach o systemie edukacji regularnie pojawiają się rozważania dotyczące tego, z czego wynikał sukces Finlandii i odniesienia do najróżniejszych statystyk dotyczących tego kraju. Wiele danych instytucji unijnych dostępnych jest na poziomie regionalnym (np. poprzez projekt ESPON – Europejską Sieć Obserwacyjną Rozwoju Terytorialnego i Spójności Terytorialnej), umożliwia więc wykonywanie analogicznej operacji nie tylko dla całych państw, ale również dla mniejszych jednostek terytorialnych.

### **5.3. Internetowe banki danych statystycznych**

Jak pisaliśmy powyżej, droga do niektórych danych publicznych jest długa i wymaga czasu oraz pracy (a bywa, że także pieniędzy). Nie oznacza to jednak, że nie można ich zdobyć. Taki wysiłek należy po prostu wpisać w plan pracy nad stanowiskiem rzeczniczym. Trzeba jednak podkreślić, że w wielu przypadkach w zupełności mogą nam wystarczyć dane dostępne w zbiorach internetowych. Łatwo się w nich zagubić – dlatego niniejsza część podręcznika ma stanowić krótki przewodnik po różnych źródłach internetowych, które mogą zawierać dane użyteczne dla działalności rzeczniczej.

Źródła te można podzielić na kilka grup. Pierwsza to serwisy internetowe publicznych instytucji. Na poziomie polskim szczegółowo opisujemy zawartość strony GUS, na poziomie międzynarodowym – zasoby Eurostatu, OECD oraz Banku Światowego. Trzeba przy tym od razu zaznaczyć, że niestety żadna z organizacji międzynarodowych nie prowadzi serwisu w języku polskim. Wszystkie są dostępne w językach angielskim i francuskim, Eurostat również w niemieckim, a strona Banku Światowego w kilkunastu różnych językach.

Kolejną grupą źródeł są pozarządowe serwisy gromadzące w jednym miejscu dane publiczne normalnie niedostępne lub dostępne w sposób rozproszony i mało przyjazny dla przeciętnego użytkownika. W naszym podręczniku opisujemy dwa z nich: mojapolis.pl, przedstawiający dane o jednostkach samorządu terytorialnego: województwach, powiatach, gminach, oraz naszakasa.org.pl, gdzie szczegółowo przedstawione są budżety wszystkich gmin w Polsce. Są to portale prowadzone przez Stowarzyszenie Klon/Jawor w ramach projektu „Dane, debata, demokracja (3D)”, którego częścią jest również niniejszy podręcznik.

Ostatnia grupa to serwisy gromadzące wyniki badań opinii. Opisujemy, gdzie możemy znaleźć tego typu dane dotyczące Polski, oraz jaki można z nich zrobić użytek.

### 5.3.1. Główny Urząd Statystyczny

Zasoby na stronie internetowej Głównego Urzędu Statystycznego ([stat.gov.pl](http://stat.gov.pl)) można zasadniczo podzielić na dwie kategorie – te zawarte w publikacjach oraz te udostępnione w interaktywnych bazach danych. W dalszej części podręcznika opisujemy dwa bazy najczęściej wykorzystywane i najbardziej przyjazne w użytkowaniu: Bank Danych Lokalnych oraz bazę Demografia. Jako że pojęcia wykorzystywane przez GUS nie muszą być zawsze oczywiste i jednoznaczne, na stronie [stat.gov.pl](http://stat.gov.pl) istnieje również dział „Metainformacje” w prawej części strony. Można tam znaleźć między innymi słownik pojęć<sup>9</sup> oraz opis wskaźników<sup>10</sup>, pomagające interpretować zgromadzone dane.

#### Publikacje

Podstawowymi wydawnictwami GUS są Roczniki Statystyczne<sup>11</sup>. Zawierają one najważniejsze wskaźniki dotyczące Polski (Rocznik Statystyczny RP) i różnych dziedzin – jest np. Rocznik

---

9 [http://www.stat.gov.pl/gus/definicje\\_PLK\\_HTML.htm](http://www.stat.gov.pl/gus/definicje_PLK_HTML.htm)

10 [http://www.stat.gov.pl/gus/opis\\_wskaznikow\\_PLK\\_HTML.htm](http://www.stat.gov.pl/gus/opis_wskaznikow_PLK_HTML.htm)

11 [http://www.stat.gov.pl/gus/roczniki\\_PLK\\_HTML.htm](http://www.stat.gov.pl/gus/roczniki_PLK_HTML.htm)

Statystyczny Przemysłu czy Rocznik Demograficzny. Jeśli chcemy zapoznać się z bardziej szczegółowymi informacjami, możemy sięgnąć do opracowań dziedzinowych, które pogrupowane są w menu „Tematy” w lewej części strony internetowej. Mogą to być notatki informacyjne, zawierające kilku czy kilkunastostronicowe podsumowania danych dotyczących określonego tematu na poziomie kraju lub wybranego regionu. Przykładem jest notatka „Ochrona Środowiska 2012”<sup>12</sup>.

Tego typu dokumenty zawierają jednak wyłącznie część danych – te, które autorzy uznali za ważne. Bardziej kompletne informacje uzyskamy z raportów z serii „Informacje i opracowania statystyczne”, takich jak np. „Oświata i wychowanie w roku szkolnym 2011/2012”<sup>13</sup>. W tym wypadku otrzymujemy szczegółowe informacje dotyczące metodologii zbierania danych oraz używanych w raporcie pojęć, rozbudowany komentarz analityków GUS, a także wiele tablic przeglądowych, w których sami możemy wyszukiwać wielkości interesujących nas wskaźników.

Na stronie GUS mamy do dyspozycji również wiele wydawnictw z zakresu statystyki regionalnej. Do większości z nich możemy dotrzeć za pośrednictwem stron urzędów statystycznych w poszczególnych województwach. Specyficznym typem publikacji jest „Statystyczne Vademecum Samorządowca” – kilkunastostronicowe portrety wszystkich jednostek terytorialnych: województw, powiatów i gmin, które przedstawiają podstawowe statystyki dla danego terytorium<sup>14</sup>. Czasem mogą one wystarczyć, jednak z reguły dobrze jest pogłębić ten obraz, sięgając do bardziej szczegółowych danych z Banku Danych Lokalnych. **[patrz obraz 5.1, str. 46]**

## Bazy danych

Bardzo często może dojść do sytuacji, w której obraz przedstawiony w publikacjach GUS wydaje się niepełny i chcemy sięgnąć do innych danych dotyczących tej samej kwestii. Do tego służą interaktywne bazy, takie jak Demografia oraz Bank Danych Lokalnych. Właśnie interaktywność jest ich kluczową cechą. Możemy swobodnie generować własne zestawienia danych, dobierając interesujące nas wskaźniki, obszar terytorialny, zakres czasowy, a w niektórych przypadkach również dodatkowe kryteria podziału (np. miasto – wieś).

Obie bazy są dość intuicyjne w obsłudze. Przy korzystaniu z nich ważne jest kilka informacji, które opisujemy poniżej. Warto również sięgnąć po metadane opisujące zawartość każdej

---

12 [http://www.stat.gov.pl/cps/rde/xbcr/gus/se\\_ochrona\\_srodowiska\\_2012\\_notatka.pdf](http://www.stat.gov.pl/cps/rde/xbcr/gus/se_ochrona_srodowiska_2012_notatka.pdf)

13 [http://www.stat.gov.pl/cps/rde/xbcr/gus/e\\_oswiata\\_i\\_wychowanie\\_2011-2012.pdf](http://www.stat.gov.pl/cps/rde/xbcr/gus/e_oswiata_i_wychowanie_2011-2012.pdf)

14 [http://www.stat.gov.pl/gus/11191\\_PLK\\_HTML.htm](http://www.stat.gov.pl/gus/11191_PLK_HTML.htm)

Strona główna | Kontakt | Mapa serwisu | BIP | RSS | English | Polityka prywatności | Szukaj

**GUS** Główny Urząd Statystyczny  
PORTAL INFORMACYJNY

Dla mediów | Kalendarium | Tajemnica statystyczna | Publikacje A-Z | **Roczniki Statystyczne 1**

Statystyczne notatki 2013

Komunikaty i obwieszczenia

- Wyniki wstępne
- Wskaźniki makroekonomiczne
- Wskaźniki monitorujące
- Sytuacja społeczno-gospodarcza Polski

Konjunktura

**BANKI I BAZY DANYCH**

- Bank Danych Lokalnych **2**
- Statystyczne Vademecum Samorządowca **3**
- Handel zagraniczny
- Demografia **4**
- SUS

**REJESTRY**

- REGION
- TERYT

**TEMATY**

- Ceny. Handel
- Ludność
- Praca. Wynagrodzenia **5**

**Uroczystość 95-lecia GUS w Pałacu Prezydenckim**

11 lipca 2013 r. (czwartek) o godz. 12.30 w Pałacu Prezydenckim odbyła się uroczystość wręczenia odznaczeń państwowych z okazji 95. rocznicy powstania Głównego Urzędu Statystycznego oraz konferencja pt.: „Jakość życia w Polsce. Aktualny stan i wyzwania w świetle badań”.  
czytaj dalej »

**Aktualności** zobacz wcześniejsze »

**Jakość życia. Kapitał społeczny, ubóstwo i wykluczenie społeczne w Polsce**  
2013-07-12

Niniejsza publikacja „Jakość życia. Kapitał społeczny, ubóstwo i wykluczenie społeczne w Polsce” zawiera analizę wyników zrealizowanego przez Główny Urząd Statystyczny (GUS) w 2011 roku nowego, wieloaspektowego i kompleksowego badania ankietowego gospodarstw domowych zabytowanego: „Badanie spójności społecznej”. Realizacja tego badania wpisuje się w szereg krajowych i międzynarodowych inicjatyw, mających na celu wzbogacenie zasobów informacji statystycznych, pozwalających na rzetelną diagnozę różnych aspektów jakości życia, która jest zalecana m.in. w Raporcie Stiglitz-Sena-Fitoussi’ego.  
czytaj dalej »

**Obroty handlu zagranicznego ogółem i według krajów I-V 2013 r.**  
2013-07-11

W okresie styczeń - maj br. eksport w cenach bieżących wyniósł 254,7 mld zł, a import 257,1 mld zł. W porównaniu z analogicznym okresem ub. roku eksport zwiększył się o 3,2%, a import zmniejszył się o 4,8%. Ujemne saldo ukształtowało się na poziomie 2,4 mld zł (w analogicznym okresie ub. roku minus 23,1 mld zł).  
czytaj dalej »

**Działalność przedsiębiorstw pośrednictwa kredytowego w 2012 roku**  
2013-07-11

Badaniem zostało objętych 69 przedsiębiorstw pośrednictwa kredytowego prowadzących działalność w 2012 r. Wśród nich było 20 spółek akcyjnych, 36 spółek z ograniczoną odpowiedzialnością, pozostałe 13 firm to inne formy spółek.  
czytaj dalej »

**Statystycznie mówiąc...**  
2013-07-11

**SPISY POWSZECHNE**

- NSP 2011 WYNIKI
- PSR 2010 WYNIKI

**Co wiesz o statystyce? Zapraszamy do zabawy Kliknij by przejść dalej...**

**MIEDZYNARODOWY ROK STATYSTYKI ORGANIZACJA UCZESTNICZĄCA**

**SPRAWOZDAWCZOŚĆ**

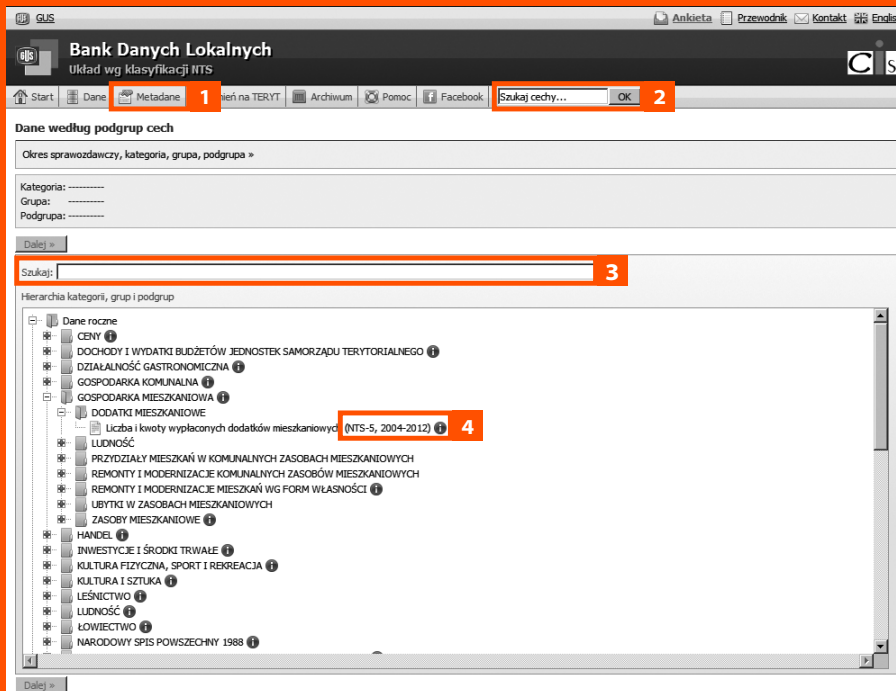
- PORTAL SPRAWOZDAWCZY **6**
- Formularze
- Badania ankietowe **7**
- Intrastat

**METAINFORMACJE** **8**

- Słownik pojęć
- Opis wskaźników (GUS)
- Opis wskaźników (UE)
- Szeregi czasowe
- Jakość w statystyce
- Klasyfikacje

Obraz. 5.1. Główna strona serwisu internetowego Głównego Urzędu Statystycznego.

1. Roczniki statystyczne; 2. Linki do Banku Danych Lokalnych; 3. Link do Statystycznego Vademecum Samorządowca 4. Linki do bazy demografia; 5. Publikacja GUS uprządkowane tematycznie; 6. Wykaz formularzy sprawozdawczych GUS; 7. Opis badań realizowanych przez GUS; 8. Metadane, czyli opis wykorzystywanych terminów i kategorii



**Obraz. 5.2. Strona Banku Danych Lokalnych GUS.**

1. Dział „Metadane”, zawierający wszystkie informacje o danych dostępnych w BDL; 2. Wyszukiwanie wszystkich wskaźników zawierających dane wyrażenie; 3. Wyszukiwanie konkretnych grup wskaźników; 4. Poziom agregacji i zakres czasowy danej grupy wskaźników.

z baz. Przede wszystkim jednak potrzebna jest praktyka – tylko spędzając trochę czasu na samodzielnym przeszukiwaniu baz, nauczymy się, gdzie kliknąć, żeby uzyskać te dane, których potrzebujemy, i będziemy wiedzieć, co one dokładnie oznaczają.

Bank Danych Lokalnych (BDL)<sup>15</sup> jest podstawowym źródłem wiedzy o jednostkach terytorialnych w Polsce. Zawiera dane z najróżniejszych dziedzin, od ludności, przez rolnictwo i rynek pracy, po sektor non-profit. Informacje w BDL możemy wyszukiwać poprzez dwie ścieżki – wychodząc od interesujących nas wskaźników (*Wg podgrup cech*) lub od terytorium, o którym chcemy się czegoś dowiedzieć (*Dla jednostki podziału terytorialnego*).

Warto zatrzymać się na chwilę przy tej drugiej ścieżce, bowiem znajomość podziału terytorialnego wykorzystywanego w statystyce publicznej przydaje się przy szukaniu i analizowaniu danych. W BDL znajdują się dane zgodne z europejską klasyfikacją NTS, w której jest pięć poziomów podziału: regiony (1), województwa (2), podregiony (3), powiaty (4) oraz gminy (5). Każda jednostka terytorialna ma swój unikalny kod, który pozwala ją zidentyfikować – kody te są zapisane w rejestrze TERYT<sup>16</sup>. Każdy wskaźnik w BDL jest dostępny na określonym poziomie agregacji – tzn. można uzyskać informacje o jego wielkości w jednostkach określonego poziomu podziału.

Jeśli interesuje nas określony wskaźnik, to na głównej stronie BDL mamy również do dyspozycji rozwijaną listę kategorii. Każda z nich dzieli się na podkategorie, w ramach których pogrupowane są poszczególne wskaźniki. Przy każdym wskaźniku na liście znajduje się informacja, na jakim poziomie agregacji dostępne są dane oraz jakiego zakresu czasowego dotyczą. **[patrz obraz 5.2, str. 47]**

### **Przykładowo:**

Wskaźnik „Liczba i kwoty wypłaconych dodatków mieszkaniowych” (Kategoria „Gospodarka mieszkaniowa”, podkategoria „Dodatki mieszkaniowe” ) posiada adnotację „NTS-5, 2004-2012”, co oznacza, że dane obejmują lata od 2004 do 2012, a ich szczegółowość sięga poziomu poszczególnych gmin.

Z kolei wskaźnik „Nakłady na badania i rozwój” (Kategoria „Nauka i technika”, podkategoria „Działalność badawczo-rozwojowa” ) ma adnotację „NTS-2, 2002-2012”, a więc możemy po-

---

<sup>15</sup> [http://www.stat.gov.pl/bdl/app/strona.html?p\\_name=indeks](http://www.stat.gov.pl/bdl/app/strona.html?p_name=indeks)

<sup>16</sup> [http://www.stat.gov.pl/bip/36\\_PLK\\_HTML.htm](http://www.stat.gov.pl/bip/36_PLK_HTML.htm)



znać jego wielkość tylko dla całego kraju, regionów lub województw. Za to zakres czasowy jest szerszy, bo obejmuje również lata 2002-2004.

Niektóre dane dostępne są na poziomie jeszcze niższym niż NTS-5 – dla poszczególnych miejscowości wchodzących w skład gminy. Wtedy przy wskaźniku pojawia się adnotacja „Miejsc”. Szczegółowy opis systemów klasyfikacji terytorialnej dostępny jest w dziale „metadane” BDL<sup>17</sup>.

Bardzo ważną zaletą Banku Danych Lokalnych jest możliwość eksportowania wybranych danych do plików XLS i CSV, które można potem edytować w arkuszach kalkulacyjnych, takich jak Microsoft Excel czy OpenOffice Calc. Nie warto za to korzystać z wykresów tworzonych przez BDL – nie spełniają one standardów, o których piszemy w kolejnym rozdziale podręcznika [**patrz podrozdział 8.3. „Zasady wizualizacji”, str. 107**].

W bazie Demografia<sup>18</sup> nie możemy wygenerować interesującego nas zestawienia danych do formatu edytowalnego. Jednak dział „Tablice predefiniowane” zawiera wszystkie dane w postaci plików XLS, musimy je tylko wyszukać według listy. Z kolei generowanie własnych zestawień jest dużo bardziej intuicyjne niż w przypadku BDL. Dane są tutaj dostępne dla jednostek podziału administracyjnego – województw, powiatów, gmin, bez poziomu regionów i podregionów.

### 5.3.2. EUROSTAT

Aby porównać Polskę z innymi krajami Unii Europejskiej, możemy skorzystać z zasobów Urzędu Statystycznego Unii, czyli Eurostatu<sup>19</sup>. Podobnie jak w przypadku GUS, również tutaj mamy do czynienia z publikacjami, w których znajdziemy dane już opracowane. Podstawowe wydawnictwa to Rocznik Eurostatu (*Eurostat Yearbook*)<sup>20</sup> oraz Rocznik Regionalny (*Regional Yearbook*)<sup>21</sup>. Poza tym do dyspozycji mamy raporty dotyczące różnych dziedzin, skonstruowane analogicznie, jak w GUS, tyle że z innym położeniem akcentów: właściwą częścią jest opracowanie analityczne wraz z wykresami, a w charakterze (obszernych) załączników znajdują się tam tablice przeglądowe oraz opis metodologii przedstawianych danych.

---

17 [http://www.stat.gov.pl/bdl/app/strona.html?p\\_name=metadane](http://www.stat.gov.pl/bdl/app/strona.html?p_name=metadane)

18 <http://demografia.stat.gov.pl/bazademografia/StartIntro.aspx>

19 <http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home>

20 [http://epp.eurostat.ec.europa.eu/statistics\\_explained/index.php/Europe\\_in\\_figures\\_-\\_Eurostat\\_yearbook](http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Europe_in_figures_-_Eurostat_yearbook)

21 [http://epp.eurostat.ec.europa.eu/statistics\\_explained/index.php/Eurostat\\_regional\\_yearbook](http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Eurostat_regional_yearbook)

Bazy danych Eurostatu są nieco trudniejsze w obsłudze niż te GUS-owskie. Nie tylko dlatego, że nie ma wersji polskiej – mają one również więcej opcji, co zwiększa ich złożoność, ale pozwala na wyciągnięcie informacji lepiej dostosowanych do naszych potrzeb. Tak samo, jak w GUS, mamy do dyspozycji „metadane”, które opisują interesujące nas wskaźniki. Część z tych materiałów została przetłumaczona na język polski.

Danych najlepiej szukać, korzystając z zakładki *Statistics Database*<sup>22</sup>, gdzie są one uporządkowane na podobnej liście rozwijanej, jak w przypadku Banku Danych Lokalnych GUS. Wybieramy kolejne dziedziny i poddziedziny, aby dotrzeć do interesujących nas tablic. Pomocna może być w tym wyszukiwarka haseł, która uwzględnia nie tylko tytuły tablic, ale również ich zawartość. **[patrz obraz 5.3, str. 51]**

Po otwarciu konkretnej tablicy mamy do dyspozycji bardzo dużo interaktywnych opcji. Gdy wejdziemy w zakładkę *Select Data*, możemy swobodnie zaznaczać, które kraje lub regiony chcemy analizować, w jakim zakresie czasowym i pod względem jakich cech z danej kategorii. Informacje z Eurostatu są jeszcze bardziej przyjazne do eksportu niż te z GUS. Mamy do dyspozycji bowiem nie tylko formaty XLS i CSV, ale również inne – np. format SAV obsługiwany przez programy do analizy statystycznej. **[patrz obrazy 5.4, str. 51 i 5.5, str. 52]**

Eurostat stwarza również duże możliwości wizualizacji danych. Dotyczy to jednak tylko części wskaźników – tych, które znajdziemy na listach *Tables by Themes* oraz *Tables on EU policy*. W tym wypadku mamy do dyspozycji wykresy oraz mapy, które obrazują interesujące nas dane. Możemy swobodnie zmieniać zakres przedstawionych danych, typ wykresu i inne opcje. Wykresy i mapy możemy wyeksportować w formie pliku PDF lub obrazu PNG. Ciekawe narzędzie wizualizacji umożliwia także serwis *Statistical Atlas*.<sup>23</sup> Jest to mapa, na której możemy przedstawić wskaźniki, które zawiera Rocznik Regionalny Eurostatu.

### 5.3.3. Organizacja Współpracy Gospodarczej i Rozwoju (OECD)

Kolejnym źródłem danych międzynarodowych – w tym wypadku w odniesieniu do 30 krajów rozwiniętych – jest strona Organizacji Współpracy Gospodarczej i Rozwoju (OECD) (<http://www.oecd.org>). OECD gromadzi wiele statystyk, które odgrywają dużą rolę w debacie publicznej. Raport „Education at a Glance” był na przykład w ostatnich latach wykorzystywany

---

22 [http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search\\_database](http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database)

23 <http://ec.europa.eu/eurostat/statistical-atlas/gis/viewer>

**Obraz 5.3. Strona internetowa Eurostatu – lista tabel ze wskaźnikami.** 1. Przeszukiwanie baz danych; 2. Metadane; 3. Wyszukiwarka wskaźników; 4. Wskaźniki przedstawione na interaktywnych tablicach; 5. Wskaźniki na interaktywnych tablicach z możliwością wizualizacji.

**Overcrowding rate by age, sex and poverty status - Total population (source: SILC)** [lic\_vho05a]

Last update: 15-07-2013  
 Interactive extraction size limit: 750000  
 Current extraction size: 360  
 Database selection: 16/76

AGE GEO INCRSP SEX TIME UNIT 1

View:  
 Setting:  Sort Ascending  Sort Descending  Sort Protocol Order  
 Show:  Codes  Labels  Both

Filtering:  
 Filtering type:  Text  Code range  Pattern  Nuts level  
 Search in:  Codes  Labels  Both

2 search Show all

Select all	Code	Label
<input checked="" type="checkbox"/>	EU28	European Union (28 countries)
<input checked="" type="checkbox"/>	EU27	European Union (27 countries)
<input checked="" type="checkbox"/>	EU15	European Union (15 countries)
<input checked="" type="checkbox"/>	NMS12	New Member States (12 countries)
<input checked="" type="checkbox"/>	EA17	Euro area (17 countries)
<input checked="" type="checkbox"/>	BE	Belgium
<input checked="" type="checkbox"/>	BG	Bulgaria
<input checked="" type="checkbox"/>	CZ	Czech Republic
<input checked="" type="checkbox"/>	DK	Denmark
<input checked="" type="checkbox"/>	DE	Germany (until 1990 former territory of the FRG)
<input checked="" type="checkbox"/>	EE	Estonia
<input checked="" type="checkbox"/>	IE	Ireland
<input checked="" type="checkbox"/>	EL	Greece

**Obraz 5.5. Widok wyboru danych wyświetlanych w tabeli.** 1. Wybór kategorii danych, które chcemy modyfikować – np. zakres czasowy lub geograficzny albo przedstawiane w tabeli wskaźniki; 2. Zaznaczanie, które dane (w tym wypadku – dla jakich krajów) mają być wyświetlane.

**OECD.StatExtracts** Complete databases available via OECD's library

Login | Version française | Contact us | User Guide | Home

Search Getting Started

Date by themes Popular queries

Find in themes 1

4

5 2 3 7

6

Study enrolled 900000: All levels of education + not allocated by level

Selection 900000: Total

Layout 900000: All educational programmes

900000: All orientation

Intensity of participation 90: Full-time and part-time

Age groups 900000: All ages

Gender 90: Total males-females

Country	Year													
	1990	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Australia	5 565 536	5 585 389	5 847 123	5 633 919	5 999 853	6 045 434	5 937 758	6 021 137	6 041 997	6 059 821	6 128 711	5 964 236	6 002 080	6 297 112
Austria	1 652 742	1 668 326	1 682 174	1 686 979	1 643 769	1 648 832	1 669 010	1 678 395	1 687 551	1 675 948	1 683 021	1 702 316	1 729 221	1 721 392
Belgium	2 608 289	2 638 301	2 704 221	2 733 590	2 774 971	2 732 428	2 792 254	2 821 180	2 829 081	2 844 592	2 868 871	2 884 291	2 907 247	2 937 247
Canada	7 057 997	7 039 833	7 047 298	7 147 318				6 486 632	(a)	0	(a)	0	(a)	0
Czech Republic	2 229 892	2 100 796	2 204 840	2 219 946	2 219 992	2 214 713	2 222 331	2 200 330	2 150 309	2 142 963	2 148 002	2 152 977	2 158 120	2 154 061
Denmark	2 215 980	1 259 919	1 259 264	1 281 245	1 295 854	1 319 853	1 377 411	1 397 896	1 394 824	1 406 558	1 402 521	1 411 190	1 426 855	1 473 470
Finland	1 219 815	1 250 158	1 278 973	1 309 645	1 318 810	1 330 292	1 345 235	1 377 334	1 388 313	1 394 548	1 398 153	1 389 187	1 398 119	1 396 562
France	14 495 237	14 415 100	14 358 552	14 292 240	14 247 289	14 359 333	14 401 428	14 939 874	14 848 248	14 880 987	14 834 661	14 805 090	14 875 582	14 925 713
Germany	16 858 908	16 913 481	16 847 078	16 913 190	16 863 525	16 842 854	16 821 686	16 699 519	16 837 284	16 670 759	16 475 459	16 370 235	16 290 875	16 301 633
Greece	2 045 371	2 001 330	2 031 340	2 052 743	2 117 624	2 102 755	2 123 181	2 194 230	2 184 995	2 107 099	2 157 590	2 182 624	2 183 041	2 178 295
Hungary	2 239 741	2 255 558	2 272 477	2 277 279	2 287 741	2 299 879	2 315 753	2 302 336	2 278 800	2 243 706	2 196 856	2 150 217	2 133 238	2 120 883
Iceland	85 731	86 351	88 361	89 151	92 514	96 354	94 069	95 803	96 191	96 944	98 856	99 651	100 981	100 364
Ireland	1 003 176	998 879	990 598	989 561	995 231	1 003 386	1 035 621	1 038 839	1 037 301	1 055 846	1 068 837	1 078 258	1 164 744	1 181 701

Legend: (a) Missing value

Data extracted on 23 Jul 2013 07:43 UTC (GMT) from OECD.Stat

Powered by Stat technology | © OECD. All rights reserved. Terms & Conditions | Privacy Policy

Login | Version française | Contact us | Site Map | User Guide | Home

**Obraz 5.6. Widok wyboru danych wyświetlanych w tabeli.** 1. Wyszukiwarka tabel; 2. Pobieranie danych w różnych formatach; 3. Tworzenie wykresów; 4. Wybór tabel według dziedzin tematycznych; 5. Zmiana układu tabeli; 6. Wybór wyświetlanych w tabeli danych; 7. Metadane – informacje o wskaźnikach.

do udowodnienia, że nauczyciele w Polsce pracują mało. Związek Nauczycielstwa Polskiego cytował go również, broniąc się przed tymi zarzutami. Pod względem działalności wydawniczej OECD przypomina zresztą bardziej międzynarodowy think tank niż urząd, taki jak GUS.

Za to, podobnie jak w przypadku GUS i Eurostatu, na stronie OECD można znaleźć pokaźne zasoby w postaci baz danych. Niestety, wiele z nich jest osiągalnych tylko w ramach płatnej subskrypcji. Dostęp do wszystkich danych dla instytucji niekomercyjnej to koszt kilku tysięcy euro (!), jeśli potrzebujemy danych z konkretnej dziedziny, cena ta i tak pozostaje na poziomie kilkuset euro. Jednak podstawowe, najważniejsze wskaźniki wykorzystywane przez OECD dostępne są w bezpłatnym serwisie StatExtracts (<http://stats.oecd.org>). Korzystanie z niego jest bardzo podobne, jak w Eurostacie. Tutaj również najpierw wybieramy interesującą nas tabelę, a później mamy możliwość modyfikacji zakresu wyświetlanych danych, tworzenia wykresów oraz eksportu do plików XLS, CSV i innych. **[patrz obraz 5.6, str. 52]**

Jeśli obsługa interaktywnych tabel nie jest dla nas dogodna, możemy do danych OECD dotrzeć również inną drogą, bezpośrednio przez główną stronę organizacji (<http://www.oecd.org/statistics>). Można stamtąd od razu ściągnąć arkusze kalkulacyjne zawierające dane odnoszące się do określonego wskaźnika.

Ciekawym serwisem związanym z posiadanymi przez OECD danymi jest jeszcze DataLab (<http://www.oecd.org/statistics/datalab>). Jest to zbiór różnego rodzaju interaktywnych narzędzi przygotowanych specjalnie do wizualizacji wybranych wskaźników. Dużą popularnością cieszył się ostatnio np. Better Life Index<sup>24</sup> – wykres pokazujący, które kraje są dla nas najbardziej przyjazne, zależnie od tego, które sfery życia są dla nas najważniejsze. Jedną z najczęściej przywoływanych publikacji jest przygotowywane regularnie tzw. Country Outlook opisujący corocznie kondycje poszczególnych krajów członków OECD.

#### 5.3.4. Bank Światowy

Jeszcze jedną organizacją międzynarodową, w której zbiory danych odgrywają ważną rolę, jest Bank Światowy. Ta instytucja otwarcie deklaruje przywiązanie do idei Otwartych Danych (Open Data). Udostępnia zgromadzoną wiedzę, podobnie jak Eurostat i OECD, poprzez serwis DataBank (<http://databank.worldbank.org/data>), gdzie znajdziemy interaktywne tablice, w których możemy dowolnie wybierać kraje, zakres czasowy i poszukiwane wskaźniki.

---

24 <http://www.oecdbetterlifeindex.org>

Możemy również tworzyć mapy i wykresy oraz eksportować dane do plików XLS i CSV.

Jednak podejście Banku Światowego do ujawniania danych idzie dalej. Po pierwsze, udostępnia on swoje zasoby nie tylko poprzez przeglądarkę internetową, ale także za pośrednictwem tzw. API (z ang. *Application Programming Interface*) – a więc oprogramowania, które robi to w sposób zautomatyzowany. Jeśli jesteś programistą (albo z nim współpracujesz), możesz dzięki API „połączyć” się na stałe do danych i np. pokazywać je na swojej stronie internetowej bez konieczności ich przeglądania i ręcznego ściągania<sup>25</sup>. Po drugie, na stronie Banku Światowego jest rozbudowany dział zawierający tzw. mikrodane (*microdata*), czyli surowe dane w formacie dostosowanym do programów statystycznych (<http://microdata.worldbank.org/index.php/home>). Część tego typu baz można ściągnąć bezpośrednio ze strony BŚ, po zalogowaniu lub nawet bez niego. W innych przypadkach podane są szczegółowe informacje, która instytucja udostępnia te dane i jak można się z nią skontaktować (z reguły są to krajowe urzędy statystyczne).

### 5.3.5. Mojapolis.pl

Mojapolis.pl to serwis internetowy prowadzony przez Stowarzyszenie Klon/Jawor, stanowiący bazę danych dotyczących społeczno-ekonomicznej kondycji województw, podregionów, powiatów oraz gmin. Informacje publikowane w [mojapolis.pl](http://mojapolis.pl) pochodzą z zasobów statystyki publicznej, danych administracyjnych oraz źródeł prywatnych i odnoszą się zarówno do cech społeczno-demograficznych mieszkańców społeczności lokalnych, jak i jakości życia, gospodarki czy środowiska naturalnego danego regionu, a także działań podejmowanych przez samorządy.

Wszystkie dane publikowane w serwisie można wyświetlić w postaci map, wykresów oraz przejrzystych tabel. [Mojapolis.pl](http://mojapolis.pl) umożliwia nie tylko zapoznanie się z wartościami poszczególnych wskaźników, lecz także generowanie porównań między wybranymi terytoriami lub zestawienie dwóch wybranych wskaźników, a także śledzenie zmian w czasie. Użytkownicy serwisu mogą samodzielnie tworzyć interesujące ich kombinacje, dowolnie zmieniać ustawienia map i wykresów.

Serwis bazuje na danych pozyskiwanych przez Stowarzyszenie Klon/Jawor. W przypadku instytucji (w tym organizacji) dysponujących wiarygodnymi i istotnymi dla opinii publicznej danymi

---

25 Więcej o tym por. <http://data.worldbank.org/node/9>

dotyczącymi poszczególnych jednostek terytorialnych możliwe jest umieszczenie ich w serwisie. Ze względu na technicznie złożony proces takiej operacji, dokonać tego może Stowarzyszenie Klon/Jawor w porozumieniu z dostawcą. W przypadku organizacji, które dostarczają danych w sposób ciągły (swoistych „hurtowników”) możliwe jest samodzielne umieszczanie nowych informacji (wymaga to jednak szkolenia i porozumienia ze Stowarzyszeniem).

Z serwisu bezpłatnie korzystać może każdy: od władz lokalnych i decydentów odpowiedzialnych za tworzenie i ewaluację polityk publicznych przez działaczy organizacji pozarządowych, aktywnych obywateli, akademików i studentów po przedstawicieli mediów. Potencjalne zastosowania są bardzo zróżnicowane, serwis pozwala m.in. na integrację i porównywanie danych z różnych źródeł, przygotowywanie lokalnych i regionalnych diagnoz czy ewaluacji polityk publicznych, tworzenie programów edukacyjnych lub rzeczniczych.

### **Jakie dane można znaleźć w serwisie?**

Dane dostępne na [mojapolis.pl](http://mojapolis.pl) pochodzą z kilkunastu różnorodnych źródeł:

- z Ministerstw: Edukacji Narodowej, Finansów, Kultury, Rozwoju Regionalnego, Rolnictwa i Rozwoju Wsi, Pracy i Polityki Społecznej, Zdrowia, Sprawiedliwości;
- z innych instytucji centralnych: Głównego Urzędu Statystycznego, Państwowej Agencji Rozwiązywania Problemów Alkoholowych, Komendy Głównej Policji, PFRON, Urzędu Komunikacji Elektronicznej;
- z instytucji regionalnych i lokalnych – np. Okręgowych Komisji Egzaminacyjnych, poszczególnych jednostek samorządu terytorialnego;
- ze źródeł prywatnych, np. udostępnionych przez Millward Brown SMG/KRC regionalnych statystyk z części prowadzonych badań.

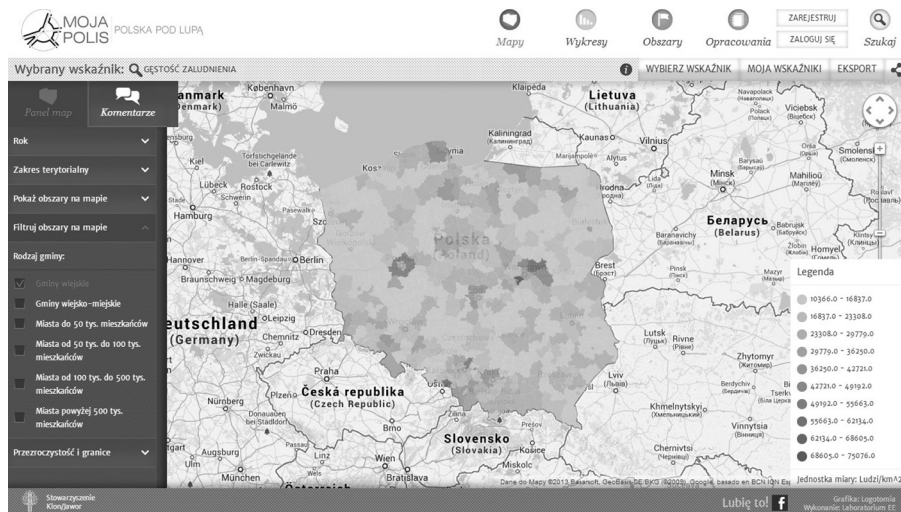
Na bazie pozyskanych danych przygotowano w serwisie ponad 1000 wskaźników, pogrupowanych w działy tematyczne, np. o mieszkańcach (dane demograficzne: liczba i struktura ludności według wieku i płci, migracje itp.), o aktywności obywatelskiej (wskaźniki aktywności społecznej, dane dotyczące wyborów oraz organizacji społecznych), o samorządach (dane o finansach samorządów i ich współpracy z organizacjami pozarządowymi), o gospodarce (wskaźniki dotyczące koniunktury gospodarczej, działań gospodarki oraz rynku pracy), o dystrybucji funduszy UE w Polsce (wartości projektów dofinansowanych z UE pogrupowanych według tematyki i programów operacyjnych), o środowisku naturalnym i jego ochronie (wskaźniki dotyczące zanieczyszczenia środowiska i obszarów chronionych), o zagospodarowaniu przestrzennym, o edukacji (wskaźniki wyposażenia szkół, wyników egzaminów, a także dostępu do edukacji przedszkolnej),

o kulturze (jej finansach, infrastrukturze i uczestnikach), o zdrowiu (dane o infrastrukturze i personelu służby zdrowia, transplantacjach, a także statystyki zachorowań) czy o jakości życia (dochody i koszty życia, warunki życia, w tym poziom bezpieczeństwa i problemy społeczne).

## Sposoby wizualizacji danych

W serwisie mojapolis.pl dostępnych jest pięć różnych typów wizualizacji danych. Każdy z nich pozwala spojrzeć na dostępne w serwisie wskaźniki z trochę innej perspektywy, a zestawienie wszystkich pięciu wizualizacji zdecydowanie ułatwia interpretację danych.

**MAPY.** Pozwalają zobaczyć zróżnicowanie Polski na poziomie województw, podregionów, powiatów czy gmin. Dzięki takiemu spojrzeniu „z lotu ptaka” możliwe jest dostrzeżenie różnic między poszczególnymi terytoriami.



Obraz 5.7. Przeglądanie map na mojapolis.pl<sup>26</sup>

26 W momencie publikacji tego podręcznika Stowarzyszenie Klon/Jawor przygotowuje się do uruchomienia nowej odsłony serwisu mojapolis.pl (planowana data uruchomienia – przełom października i listopada 2013 r.). Wszystkie zamieszczone niżej grafiki serwisu oraz opis jego funkcjonalności odnoszą się już do tej nowej wersji. W związku z ciągle toczącymi się pracami nad jego ulepszeniem, niektóre z zamieszczonych grafik i opisów mogą nie oddawać ostatecznego kształtu serwisu.



Korzystając z map, użytkownik serwisu może samodzielnie wybrać interesujący go poziom agregacji danych i rok. Możliwe jest przybliżanie (zoom) i zaznaczanie konturem wybranych obszarów. Dostępnych jest także kilka rodzajów filtrów (w zależności od wskaźnika): według typów obszarów – pozwalający oddzielnie oglądać obszary wiejskie, duże i małe miasta, według płci – z możliwością wyboru danych tylko dla kobiet lub tylko dla mężczyzn, a także według określonych przez użytkownika wartości wskaźnika (np. 20% najlepszych lub najgorszych).

**WYKRESY.** W serwisie dostępne są cztery rodzaje prezentacji danych:

**Wykresy rozrzutu** umożliwiają zestawienie wartości z różnych wskaźników i zaobserwowanie korelacji między nimi, np. możemy sprawdzić, czy najczęściej śmieci produkują najbogatsze gminy. Możliwość zestawienia dwóch wskaźników pozwala na formułowanie wielu ciekawych hipotez na temat współwystępowania różnego rodzaju zjawisk, czego często najbardziej nam brakuje w suchych, statystycznych danych. Wykresy na [mojapolis.pl](http://mojapolis.pl) są animowane – pozwalają śledzić zmiany w relacjach między dwoma wskaźnikami w czasie (por. obraz 5.7).

**Wykresy liniowe zmian w czasie.** Na tym wykresie możemy obejrzeć zmiany w czasie wartości danego wskaźnika dla jednego lub wielu obszarów. Dane w serwisie pozwalają na porównanie kilku, a niekiedy nawet kilkunastu lat. Patrząc na linię trendu, łatwo można odpowiedzieć na pytanie, czy sytuacja danego obszaru się zmienia, a jeśli tak, to w którym kierunku te zmiany prowadzą. Dodatkowo można porównać, jak rozwój na danym obszarze wygląda na tle podobnych obszarów – np. jak zmieniała się wartość inwestycji Wrocławia w lokalny transport zbiorowy w ostatnich latach na tle zmian w innych metropoliach.

**Wykresy słupkowe** pokazujące rozkład wartości danego wskaźnika, czyli skalę różnic między poszczególnymi województwami, podregionami, powiatami czy gminami. Pozwalają one przyjrzeć się, czy są regiony zdecydowanie wyróżniające się na tle innych (na plus lub minus). Łatwo je zidentyfikować – po najechaniu na poszczególne słupki kursorem można odczytać nazwę gminy, powiatu lub województwa, a także wartość wskaźnika.

**Tabele rankingowe,** w których dane dla województw, podregionów, powiatów lub gmin zostały uporządkowane ze względu na wartość wskaźnika, tworząc ranking. Z tabeli jeszcze łatwiej niż na wykresie słupkowym można odczytać, które obszary zajmują pierwsze i ostatnie miejsca w Polsce pod względem wartości danego wskaźnika. Możemy np. zobaczyć, na którym miejscu plasuje się nasze województwo czy powiat pod względem poziomu przestępczości.



**Obraz 5.8. Wykres rozrzutu na mojanpolis.pl**

### Co jeszcze można znaleźć w serwisie?

**OBSZARY.** Każde województwo, powiat i gmina w Polsce ma tu swoją wizytówkę, gdzie znaleźć można podstawowe dane, a także linki do ważnych stron. Dostępna jest także tabela zestawiająca dane o wydatkach samorządów gminnych ze wskaźnikami kontekstowymi, np. możemy sprawdzić, ile miasto Warszawa wydało na gimnazja i od razu dowiedzieć się, ile spośród warszawskich gimnazjów jest wyposażonych w pracownie językowe.

**MOJA POLIS** POLSKA POD LUPĄ

POLSKA - WOJEWÓDZTWO MAZOWIECKIE

Warszawa

### Podstawowe dane

Wskaźnik	Warszawa	Średnia	Zobacz mapy/wykresy	Usługi
Liczba ludności	1 720 398 os.	16 160 os.	<input type="checkbox"/>	<input type="checkbox"/>
Powierzchnia	517 km²	125 km²	<input type="checkbox"/>	<input type="checkbox"/>
Średnia wieku	41,9 lat	38,5 lat	<input type="checkbox"/>	<input type="checkbox"/>
Bezrobotni wśród ludności w wieku produkcyjnym	3,9 %	9,4 %	<input type="checkbox"/>	<input type="checkbox"/>

Dodaj inny wskaźnik

### Finanse gminy Warszawa i ich kontekst

Wskaźnik finansowy	Wskaźnik kontekstowy
<b>Edukacja</b>	
Wydatki gmin - gimnazja	% gimnazjów wyposażonych w pracownię językowe
311 916 tys.	34%
	% gimnazjów wyposażonych w pracownię komputerowe
	71%
	Wyniki egzaminów gimnazjalnych od 2012 r. - język polski
	73%
Wydatki gmin - LO	Wyniki matur 2012 r. - język polski
310 053 tys.	92%
<b>Transport i kultura</b>	

**Obraz 5.9. Wizytówki obszarów na mojapolis.pl**

**OPRACOWANIA.** Tu znajdziemy przygotowane przez zespół mojapolis.pl opracowania i analizy, które ułatwiają interpretację publikowanych w serwisie danych i mogą być cennym źródłem informacji np. dla dziennikarzy czy studentów. Teksty można przeszukiwać według kategorii, tagów, a także sortować według daty dodania, autora czy liczby polubień.

Dostępne w serwisie dane można eksportować w formacie edytowalnym (xls, csv), a wizualizacje i opracowania – w formacie pdf.

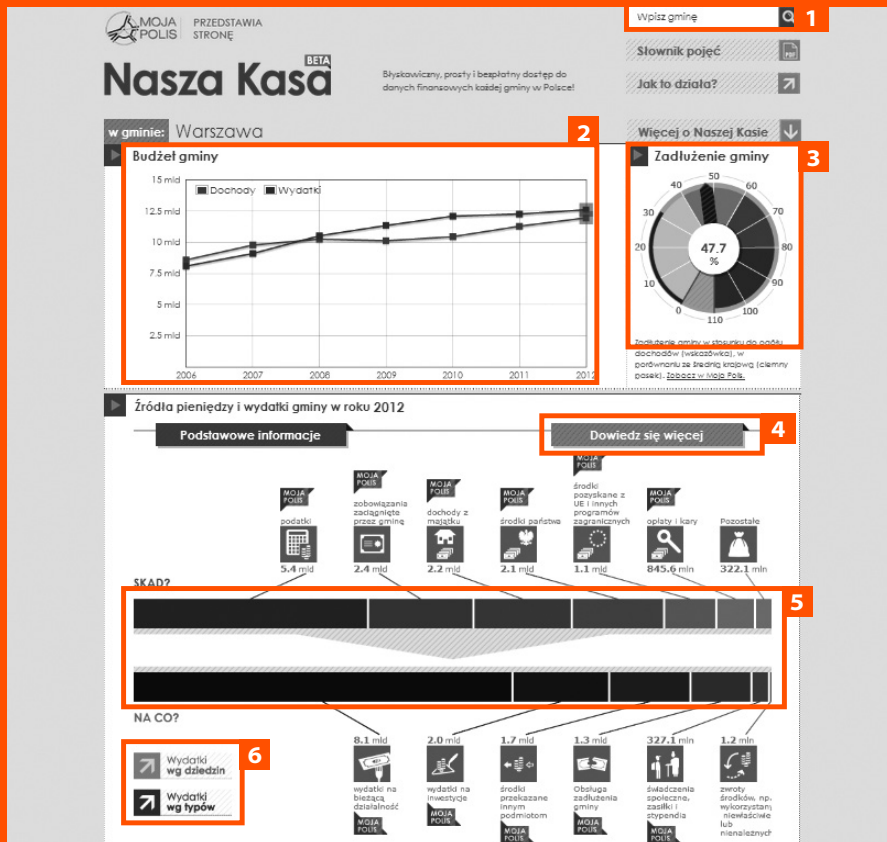
## Porównanie baz danych udostępnianych przez instytucje opisane powyżej

	Interaktywne tworzenie tabel	Eksport danych (format plików)	Wizualizacja danych	Metadane (opis wskaźników i metodologii)	Dostęp do mikro danych (danych jednostkowych)
Bank Danych Lokalnych GUS	+	CSV, XLS	Mało użyteczne wykresy	+	-
Baza Demografia GUS	+	-	-	+	-
Eurostat	+	CSV, XLS, HTML, PX, SAV, TSV, PDF	Interaktywne mapy i wykresy dla niektórych danych	+	Ograniczony, wyłącznie do celów naukowych
OECD StatExtracts	+	CSV, XLS, PX, XML	Interaktywne mapy i wykresy	+	-
Bank Światowy	+	CSV, XLS, XML	Interaktywne mapy i wykresy	+	Częściowy
Mojapolis.pl	+	CSV, XLS, PDF	Interaktywne mapy i wykresy	+	-

### 5.3.6. Naszakasa.org.pl

Nszakasa.org.pl to serwis uruchomiony w 2012 roku przez Stowarzyszenie Klon/Jawor, jako przedsięwzięcie satelitarne wobec mojapolis.pl, opisanego w poprzednim podrozdziale.

Jest to interaktywna przeglądarka informacji o finansach jednostek samorządu terytorialnego szczebla gminnego. Daje nam bezpłatny dostęp do szczegółowych danych na temat dochodów i wydatków wszystkich gmin, a także podstawowych informacji o ich kondycji finansowej (zadłużeniu) począwszy od 2006 r. Obecnie znajdują się w nim dane aż do 2012 r., które będą sukcesywnie uzupełniane.



**Obraz 5.10. Widok dla wybranej kategorii wydatków lub dochodów gminy**

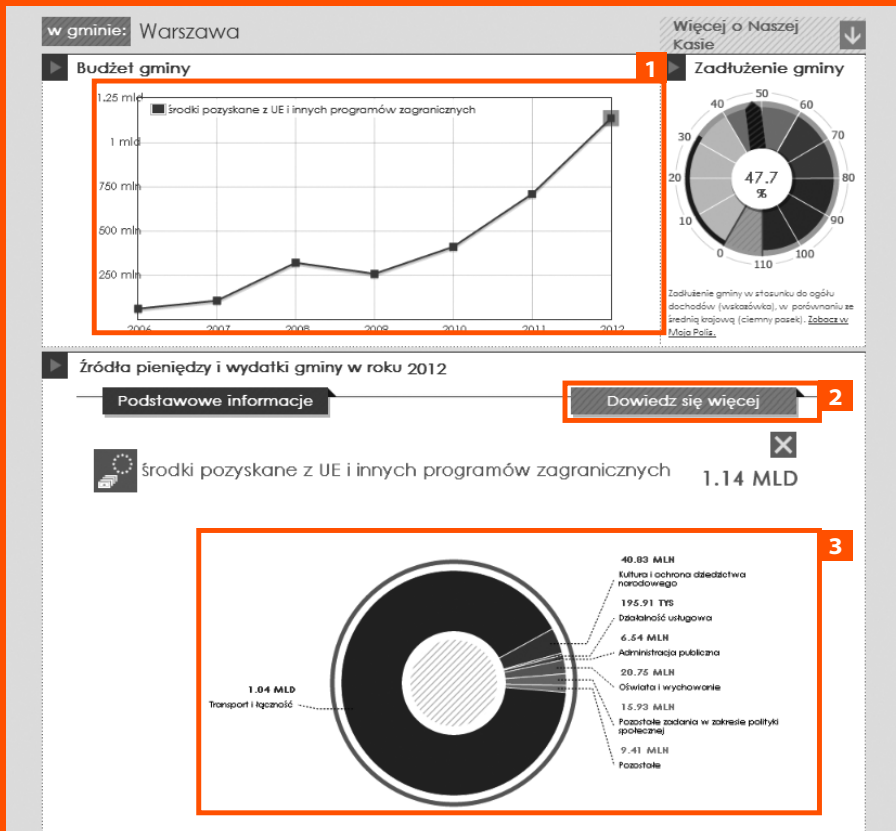
1. Wybór gminy/miasta; 2. Dynamika dochodów i wydatków gminy; 3. Wskaźnik zadłużenia gminy; 4. Zaawansowany widok struktury budżetu; 5. Struktura budżetu; 6. Zmiana wyświetlanego sposobu klasyfikacji wydatków.

Naszakasa.org.pl została zaprojektowana jako rodzaj mozaiki złożonej z kilku wykresów, przedstawiających dane o sytuacji finansowej danego samorządu. Znajdziemy tu wykres liniowy ukazujący dynamikę przychodów i wydatków gminy, wskaźnik jej zadłużenia, a niżej dwie belki obrazujące najistotniejsze typy dochodów i wydatków samorządowych, opisane etykietami oraz za pomocą ikonek. Ponad wykresami znajduje się okno wyszukiwarki umożliwiającej wybór interesującej użytkownika gminy (można też uzyskać zbiorcze dane o kondycji wszystkich gmin w Polsce, wpisując w wyszukiwarkę hasło „Polska”). **[patrz obraz 5.9, str. 61]**

Przeglądarka bazuje na informacjach ze sprawozdań budżetowych składanych przez samorządy do Ministerstwa Finansów RP. Dane te agregowane są w układzie rocznym i prezentowane z zastosowaniem autorskiej, stworzonej przez twórców serwisu klasyfikacji, bazującej na rozporządzeniu Ministra Finansów w sprawie szczegółowej klasyfikacji dochodów, wydatków, przychodów i rozchodów oraz środków pochodzących ze źródeł zagranicznych. Rozporządzenie, nazywane często klasyfikacją budżetową, jest podstawą budżetowania i księgowania wydatków we wszystkich jednostkach administracji publicznej w Polsce. Niestety jest ono jednak mało użyteczne dla przeciętnego obywatela ze względu na hermetyczność stosowanego w nim języka, a także ze względu na złożoność schematu klasyfikacji.

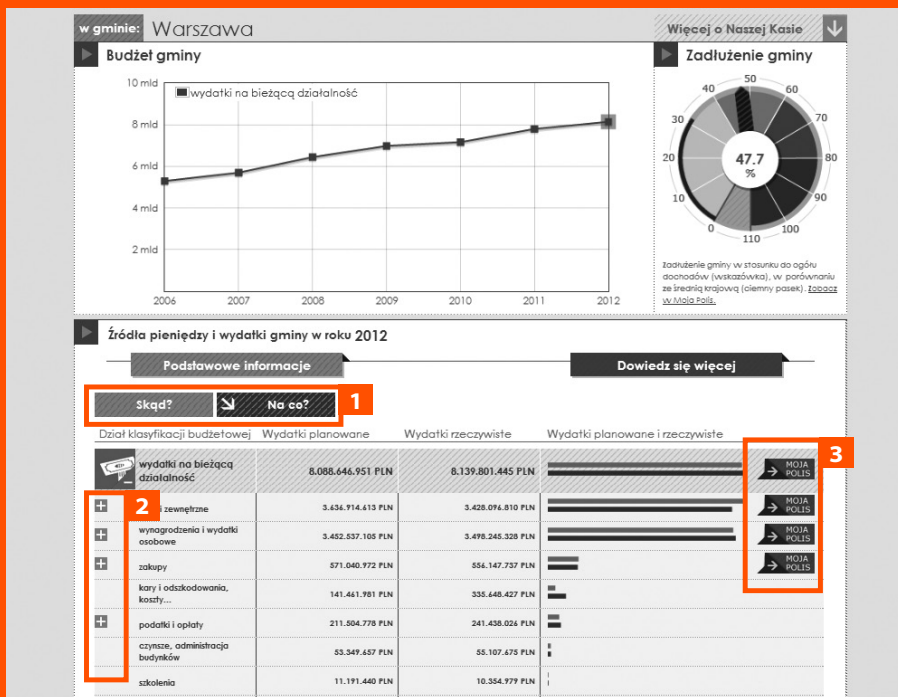
Naszakasa.org.pl grupuje poszczególne składowe klasyfikacji budżetowej w części, które są bardziej czytelne dla opinii publicznej, oraz dąży do nazywania ich prostszym, bliższym codzienności językiem. Bierze przy tym pod uwagę – na tyle, na ile to możliwe – zmiany w klasyfikacji budżetowej – tak, aby interesujące nas kategorie wydatków były porównywalne w czasie. Dokument opisujący relacje między klasyfikacją budżetową i grupowaniami serwisu oraz ograniczenia porównywalności danych w czasie można znaleźć na stronie.

Wykresy nie są jedynie sposobem na wizualizację danych, ale również narzędziami do ich szczegółowej eksploracji. Wykres zmian w czasie (liniowy) pozwala wybrać okres, z którego chcemy obejrzeć dane – aby to zrobić, wystarczy kliknąć na odpowiedni punkt na linii wydatków lub dochodów. Następnie wszystkie inne wykresy na stronie zaczynają pokazywać dane z tego właśnie roku. Kliknięcie na którykolwiek z elementów wykresu belkowego pokazującego strukturę wpływów i wydatków danej gminy powoduje pojawienie się na pierwszym planie wykresu kołowego, który obrazuje strukturę dochodów/wydatków w klikniętej kategorii. Sprawia też, że na wykresie liniowym linie dotyczące dochodów i wydatków ogółem zastępowane są linią ukazującą zmiany w czasie wybranej przez nas kategorii. **[patrz obraz 5.11, str. 63]**



**Obraz 5.11, Zaawansowany widok struktury budżetu**

1. Dynamika wybranej kategorii dochodów lub wydatków; 2. . Zaawansowany widok struktury budżetu; 3. Struktura dochodów lub wydatków wybranej kategorii;



**Obraz. 5.12. Zaawansowany widok struktury budżetu**

1. Przełączanie pomiędzy dochodami i wydatkami; 2. Rozwijanie listy podkategorii; 3. Przejście do wizualizacji różnicowników między gminami w serwisie Moja.Polis.pl



Serwis oferuje też możliwość przełączania się między dwoma różnymi sposobami klasyfikacji wydatków samorządowych. Mogą one być szeregowane według dziedzin, czyli działów klasyfikacji budżetowej (przy czym, jak była już o tym mowa, naszakasa.org.pl nieco upraszcza nazwy działów, aby były krótsze i bardziej zrozumiałe dla użytkownika). To optymalny wybór, jeśli interesują nas wydatki gminy na określone zadania – na przykład na kulturę, pomoc społeczną czy transport. Druga opcja to klasyfikowanie wydatków według typów. Pozwala ona na dotarcie do danych o określonych kategoriach kosztów gminy bez względu na to, jakiej dziedziny dotyczyły. Wybieramy ją, jeśli chcemy dotrzeć do danych na temat wydatków samorządu na wynagrodzenia, inwestycje, usługi zewnętrzne itp.

Kolejne ciekawe narzędzie to możliwość przełączania się między widokiem podstawowej struktury wpływów/wydatków samorządowych a widokiem zaawansowanym. Umożliwia nam to zakładka *dowiedz się więcej* nad wykresem belkowym. Kliknięcie w nią przenosi nas do widoku wykresu słupkowego, który umożliwia głębszą eksplorację danych o wpływach i wydatkach. Wykres jest jednocześnie rodzajem interaktywnego drzewa dochodów/wydatków: kliknięcie w jego poszczególne kategorie (słupki na wykresie) powoduje rozwinięcie tych kategorii w taki sposób, że pod interesującym nas słupkiem pojawia się osobna wizualizacja składających się na niego kategorii. Możemy w ten sposób przeglądać wydatki aż do podgałęzi czwartego rzędu, a więc do poziomu bardzo drobnych wpływów/wydatków. I tak, chcąc dowiedzieć się, jakie były wydatki miasta Łódź na cele osobowe i wynagrodzenia pracowników tamtejszych gimnazjów, musimy (po znalezieniu Łodzi przez wyszukiwarkę w prawym górnym rogu serwisu) kliknąć na belkę *dowiedz się więcej* pod wykresem liniowym, a następnie na przycisk *na co?* i *wydatki wg dziedzin*, który dotyczy wydatków tego miasta na poszczególne zadania. Na wykresie ukazują się wszystkie kategorie wydatków Łodzi – ponieważ jednak interesują nas tylko wydatki na gimnazja, klikamy w kategorię *oświata, wychowanie*. Pod klikniętą belką ukazuje się zestawienie wydatków na ten cel, z której wybieramy z kolei kategorię *gimnazja*. Wiemy już, że wydatki osobowe i na wynagrodzenia w łódzkich gimnazjach w 2012 roku wyniosły niemal 77 milionów złotych (a z wykresu liniowego dowiadujemy się, że stale rosną). Co ważne, przeglądarka podaje nam dane nie tylko na temat rzeczywistych wydatków gminy w wybranej kategorii, ale także na temat wydatków planowanych – czyli takich, które zakładano dla tej kategorii na początku danego roku finansowego. Analiza rozbieżności między tymi planami i ich realizacją może być interesującym przyczynkiem do dyskusji na temat zdolności samorządu do planowania działań. **[patrz obraz 5.12, str. 64]**

Naszakasa.org.pl została zaprojektowana przede wszystkim jako przeglądarka finansów konkretnej gminy. Jeśli jesteśmy zainteresowani porównywaniem danych między różnymi samorządami, możemy łatwo skorzystać z możliwości, jakie oferuje w tym względzie macierzysty serwis – mojanapolis.pl. Przy głównych kategoriach wpływów i wydatków danej gminy, a także przy liczniku jej zadłużenia, widnieją ikonki z napisem *mojanapolis*. Kliknięcie w którąś z nich powoduje przeniesienie do nowej karty przeglądarki internetowej, na której otwarta zostaje wizualizacja zróżnicowań związanych z daną kategorią wpływów/wydatków na mapie. Możemy następnie dowolnie analizować te różnice, korzystając ze wszystkich opcji mojanapolis.pl, o których była mowa w poprzednim podrozdziale.

### 5.3.7. Dostępne wyniki badań opinii publicznej

Jeszcze jednym źródłem danych, które często wykorzystywane są w debacie publicznej i mogą być ważne w działalności rzeczniczej, są wyniki badań opinii. W Polsce coraz częściej przeprowadzają je prywatne agencje, zrzeszone w Organizacji Firm Badania i Rynku (OFBOR). Robią to bądź na zlecenie np. gazety lub organizacji pozarządowej, bądź z nastawieniem na sprzedaż wyników. Dlatego jeśli sami nie zamawialiśmy danego badania, nie mamy raczej szans na dostęp do szczegółowych danych, chyba że dużo za nie zapłacimy – w innym wypadku pozostają nam wyłącznie te wnioski, które pojawiły się w mediach.

Wciąż jednak prowadzone są również badania publiczne. Najważniejszą instytucją odpowiedzialną za nie jest **Centrum Badania Opinii Społecznej**. Jest to państwowa fundacja, której działalność reguluje odrębna ustawa. Od początku lat 90-tych CBOS co miesiąc realizuje sondaż „Aktualne problemy i wydarzenia”, w którym poruszane są różne zagadnienia – przy czym pewne pytania zadawane są z dużą regularnością. Dlatego wyniki badań CBOS stanowią nieocenione źródło danych do porównań dotyczących ostatnich 20 lat.

Podstawowe wyniki niekomercyjnych badań CBOS dostępne są bezpłatnie na stronie fundacji ([www.cbos.pl](http://www.cbos.pl)) w formie krótkich raportów podsumowujących określone zagadnienia – jak np. „Nieufni czy zapominalscy? O odpisie podatkowym na rzecz organizacji pożytku publicznego”<sup>27</sup> czy „Aspiracje i motywacje edukacyjne Polaków w latach 1993-2009”<sup>28</sup>. Można je znaleźć poprzez listę wszystkich raportów<sup>29</sup> z podziałem na lata i miesiące bądź poprzez

---

27 <http://badanie.cbos.pl/details.asp?q=a1&id=4835>

28 <http://badanie.cbos.pl/details.asp?q=a1&id=4135>

29 <http://www.cbos.pl/PL/publikacje/raporty.php>

wyszukiwarkę raportów<sup>30</sup>. Część surowych danych jest też dostępna w Archiwum Danych Społecznych, o którym więcej piszemy poniżej.

Poza krótkimi komunikatami, których powstaje kilkanaście-kilkadziesiąt w miesiącu, od czasu do czasu CBOS wydaje również obszerniejsze raporty z serii „Opinie i diagnozy”. Są to kompleksowe opracowania gromadzące wyniki różnych badań CBOS na dany temat. Te publikacje trzeba już kupić – w formie elektronicznej lub drukowanej.

Poza CBOS badania opinii realizują także instytucje naukowe. Najbardziej rozpoznawalną markę wyrobiło sobie w Polsce badanie „**Diagnoza Społeczna**”. Realizowane jest przez niezależną Radę Monitoringu Społecznego pod przewodnictwem prof. Janusza Czapińskiego, a powstaje częściowo z pieniędzy publicznych (m.in. Ministerstwa Pracy i Polityki Społecznej), a częściowo ze środków przekazanych przez prywatnych sponsorów. Badanie odbywa się co 2 lata na dużej próbie kilkunastu tysięcy osób i ma charakter panelowy, tzn. do niektórych respondentów wraca się w kolejnych latach. Zakres tematyczny badania najlepiej wyjaśnia jego podtytuł: „Warunki i jakość życia Polaków”.

Wyniki „Diagnozy Społecznej” spotykają się z dużym oddźwiękiem i są komentowane zarówno przez media, jak i przez polityków. Co ważne z naszego punktu widzenia, każdy ma do nich dostęp i może samodzielnie przeanalizować zebrane dane. Ze strony [www.diagnoza.com](http://www.diagnoza.com) można pobrać nie tylko raporty ze wszystkich edycji badania, począwszy od 2000 roku, ale również kwestionariusze, które pozwalają zorientować się, jakie dokładnie pytania były zadawane respondentom. Znaleźć tam także można bazy danych surowych, które pozwalają na przeprowadzanie analiz statystycznych.

Liczne bazy danych przeznaczone do samodzielnej obróbki statystycznej zawiera również **Archiwum Danych Społecznych** – projekt Instytutu Filozofii i Socjologii Polskiej Akademii Nauk. Na stronie [www.ads.org.pl](http://www.ads.org.pl) znajdziemy wyniki niektórych badań CBOS, Diagnozy Społecznej, ale również wielu innych badań opinii o charakterze akademickim – m.in. Polskiego Generalnego Sondażu Społecznego i Polskiego Generalnego Studium Wyborczego. Istotne ograniczenie tego zbioru jest takie, że jest on użyteczny wyłącznie dla osób, które znają się na statystyce.

---

30 <http://badanie.cbos.pl/search.asp?srchtype=general>

Jeśli interesują nas **porównania międzynarodowe**, możemy odwołać się do kilku popularnych źródeł gromadzących dane sondażowe o takim charakterze. Pierwszym adresem może być baza tzw. Eurobarometrów – badań sondażowych realizowanych przez Komisję Europejską we wszystkich krajach Unii, w tym w Polsce ([http://ec.europa.eu/public\\_opinion/index\\_en.htm](http://ec.europa.eu/public_opinion/index_en.htm)). Uzyskamy tu łatwy dostęp m.in. do tzw. „standardowych Eurobarometrów“ (raportów z cyklicznych badań sondażowych o powtarzalnej strukturze, dotyczących m.in. opinii na temat Unii Europejskiej, jej instytucji, sytuacji ekonomicznej, problemów wskazywanych przez mieszkańców poszczególnych krajów itp.), a także Eurobarometrów „specjalnych”, poświęconych różnym specyficznym zagadnieniom. To prawdziwa kopalnia wiedzy, jednak niestety dostępna tylko w językach angielskim, francuskim i (częściowo) niemieckim.

Zainteresowani porównaniami międzynarodowymi zainteresują się też innymi globalnymi projektami sondażowymi (dostępnymi głównie w języku angielskim), takimi jak European Social Survey (<http://www.europeansocialsurvey.org>), European/World Values Survey (<http://www.worldvaluessurvey.org>) lub General Social Survey<sup>31</sup> czy Gallup World Poll<sup>32</sup>. Ich dane i/lub bazujące na nich opracowania są dostępne nieodpłatnie. Istnieją też podobne projekty dedykowane porównawczym badaniom określonych zagadnień (np. globalnych porównań jakości edukacji badanych w projekcie PISA<sup>33</sup>). Nie ma tu miejsca, aby omawiać je wszystkie – namawiamy jednak, żeby ich szukać i korzystać z nich zawsze, kiedy potrzebujemy międzynarodowego odniesienia dla naszych kampanii.

Poza wymienionymi tu serwisami, które umożliwiają dostęp do narzędzi oraz surowych wyników badań (zbiorów danych) istnieją też serwisy, które specjalizują się w „metainformacji” (można powiedzieć informacji o informacji). Ich celem jest pomoc w nawigowaniu i docieraniu do ważnych dla pewnego (często specyficznego) rodzaju działań – danych i opracowań. Jednym z nich, wymienionym tu na prawach przykładu, jest serwis Civicpedia prowadzony przez wspomniane już kilkakrotnie Stowarzyszenie Klon/Jawor ([www.civicpedia.ngo.pl](http://www.civicpedia.ngo.pl)). Zbiera on wiedzę o wynikach badań opisujących szeroko rozumianą aktywność obywatelską i działalność organizacji pozarządowych.

---

31 <http://www3.norc.unc.edu/gss+website>

32 <http://www.gallup.com/strategicconsulting/en-us/worldpoll.aspx>

33 <http://nces.ed.gov/surveys/pisa>

## 6. Przeprowadzanie i zlecenie badań

W działaniach rzeczniczych zdarza się, że istniejące, dostępne nam dane nie są wystarczające. Takie sytuacje mają miejsce zwłaszcza, gdy zajmujemy się tematem nie leżącym w obszarze zainteresowań urzędów statystycznych i instytucji badawczych, kiedy dostępne dane nie są aktualne lub chcemy zdobyć bardziej szczegółową wiedzę, np. odnoszącą się do poziomu lokalnego. Możemy też uznać, że obraz sytuacji zawarty w dostępnych źródłach jest zbyt powierzchowny i zależy nam na pogłębionej analizie jakiegoś zjawiska, np. poprzez uzyskanie wyników badań jakościowych. Zdarza się zatem, że organizacje stają przed wyzwaniem, którym jest samodzielne przeprowadzenie lub zlecenie badań. Jak się za to zabrać?

Musimy na wstępie uczynić dwa zastrzeżenia. Po pierwsze, w podręczniku poruszamy wyłącznie problematykę badań społecznych. Zdajemy sobie sprawę, że wiele zagadnień, którymi zajmują się organizacje rzecznicze, wymaga specjalistycznych badań innej natury, np. dotyczących ochrony środowiska, transportu itp. Jako autorzy podręcznika nie mamy kompetencji do tego, żeby udzielać porad w tym zakresie. To, że piszemy wyłącznie o badania społecznych, wynika także z faktu, że tego rodzaju badania mogą mieć walor uniwersalny i okazać się potrzebne każdej organizacji, również takiej, która zajmuje się „twardymi” kwestiami, kiedy chce poznać punkt widzenia, opinie mieszkańców jakiegoś terytorium lub całego kraju na temat kwestii, którymi się zajmuje.

Drugie zastrzeżenie jest takie, że nie jest to w żadnym wypadku podręcznik, który dawałby kompletny zakres wiedzy do realizacji badań społecznych, a dla osób zajmujących się regularnie badaniami uczynione tu uwagi będą dość oczywiste. Podręczników metodologicznych jest jednak sporo i na pewno zaspokoją one potrzeby zarówno osób poszukujących podstawowych informacji, jak i chcących zdobyć bardziej zaawansowaną wiedzę. Niektóre z tych książek wymieniamy w spisie polecanych lektur, który umieściliśmy na końcu książki.

W niniejszej publikacji ograniczamy się do udzielenia kilku praktycznych wskazówek, dostosowanych do specyfiki działania organizacji pozarządowych. Dotyczą one przede wszystkim zlecenia badań na zewnątrz. W tego typu okolicznościach niedoświadczona organizacja może łatwo się pogubić i wydać spore środki na produkt, który czasami okazuje się niesatysfakcjonujący czy zgoła bezużyteczny. Mamy nadzieję, że nasze uwagi pomogą zapobiegać takim sytuacjom.

## 6.1. Samodzielne przeprowadzanie badań

Zarówno w przypadku samodzielnego podjęcia się przeprowadzania badań, jak i zlecenia tego zadania na zewnątrz – firmie lub indywidualnemu badaczowi, musimy dobrze wiedzieć, co chcemy osiągnąć, a w szczególności, na jakie pytania szukamy odpowiedzi. O wadze etapu konceptualizacji badania pisaliśmy już wcześniej [patrz podrozdział 3.2. „Konceptualizacja”, str. 20], jednak jest to na tyle istotne, że trzeba to jeszcze raz podkreślić: żadna techniczna wiedza na temat metod badawczych nam nie pomoże, jeśli dobrze nie przemyślimy koncepcji. Dopiero po jasnym ustaleniu celu badań możemy decydować o metodach, włącznie z tak zasadniczym wyborem jak, czy mają to być badania ilościowe, czy jakościowe (patrz tabela poniżej).

	BADANIA ILOŚCIOWE	BADANIA JAKOŚCIOWE
Cel	Pomiar, szacowanie, badanie zależności	Zrozumienie zjawiska, procesów
Pytania badawcze	Ile? Jak dużo? Jak często?	Jak? Dlaczego? Po co? W jaki sposób?
Techniki badawcze	Wystandaryzowany kwestionariusz z przewagą zamkniętych pytań	Swobodny wywiad, obserwacja
Próba	Duża, setki lub tysiące osób; dobierana w sposób losowy lub kwotowy	Z reguły kilkanaście-kilkadziesiąt osób; dobierana w sposób celowy
Sposoby analizy i interpretacji wyników	Analiza statystyczna umożliwiająca uogólnienie wyników na całą populację	Wyniki opierające się na subiektywnej interpretacji

Organizacje często decydują się na badania za pomocą metod ilościowych. Nierzadko, szczególnie wśród osób mniej obeznanymi z warsztatem, istnieje błędne przekonanie, że tylko „procenty i liczby” są obiektywnymi i poważnymi badaniami. Nie chcemy umniejszać wagi badań ilościowych, ale pamiętajmy, że ich prowadzenie (szczególnie jeśli istotnie mają stanowić „mocne dowody”) oznacza wiele wyzwań, zarówno jeśli chodzi o przygotowanie narzędzi badawczych (np. treść pytań w ankiecie), jak i samego etapu realizacji, a zwłaszcza jego kosztów związanych z zapewnieniem reprezentatywności i wiarygodności. Chcielibyśmy podzielić się dwiema wskazówkami:

- Tworząc narzędzia badawcze, warto inspirować się tym, jak robią to inni (rzecz jasna, jeśli uważamy, że są to instytucje wiarygodne). Dobrze jest przejrzeć informacje dotyczące badań realizowanych w podobny sposób lub poświęconych zbliżonym zagadnieniom i sprawdzić

(o ile to możliwe), jakie pytania zadawano, jak wyglądał dobór respondentów itp. W przypadku niektórych większych projektów badawczych dostępna jest pełna dokumentacja metodologiczna, włącznie z kwestionariuszami [patrz podrozdział 5.3.7. „Dostępne wyniki badań opinii publicznej”, str. 66].

- Istnieją sposoby zmniejszenia kosztu samodzielnego realizowania badań. Warto rozważyć wykorzystanie bezpłatnych aplikacji komputerowych, takich jak Formularze Google czy LimeSurvey, które umożliwiają przeprowadzenie badania ankietowego przez Internet (patrz ramka). Trzeba przy tym pamiętać o ograniczeniach tego typu rozwiązań – w większości przypadków ankieta internetowa będzie mniej wiarygodna niż przeprowadzona metodą tradycyjną. Po pierwsze dlatego, że w wielu tzw. sondach można odpowiadać nieograniczoną liczbę razy. Tę trudność łatwo pokonać poprzez zastosowanie procedur zapewniających kontrolę tego, kto ją wypełnia (np. kontrola numeru IP lub indywidualne kody dostępu, tzw. *tokens*). Po drugie, w przypadku ankiety internetowej dużo łatwiej może dojść do sytuacji, w której wypełniać ją będą przypadkowe osoby, co spowoduje brak reprezentatywności badania. Jednak jeśli uda nam się skonstruować odpowiednią próbę i dotrzeć z ankietą do wybranych osób, to jej wyniki mogą być równie wiarygodne, jak w tradycyjnych badaniach.

---

## Formularze Google

Formularze Google to element pakietu Google Apps. Jest to najprostszy, bezpłatny sposób na stworzenie internetowej ankiety. Aplikację obsługuje się poprzez Dysk Google (<http://drive.google.com>). Oznacza to, że dane gromadzone są w chmurze Google, w związku z czym należy pamiętać o ich zgraniu na własny dysk lub serwer. Trzeba też pamiętać, że korzystając z tego rodzaju darmowych narzędzi, na ogół nie jesteśmy wyłącznym ich dysponentem. Dostęp do nich ma także sam dostawca narzędzia (w tym wypadku firma Google).

Podstawowe cechy Google Forms:

- kilka rodzajów pytań (jednokrotnego i wielokrotnego wyboru, skale, pytania otwarte itp.);
- możliwość wspólnej pracy nad formularzem przez wiele osób;
- przechowywanie formularza i bazy odpowiedzi na internetowym dysku (w chmurze);
- udostępnianie ankiet respondentom poprzez link do formularza lub umieszczenie ankiety na stronie internetowej;
- prosty system podglądu odpowiedzi, umożliwiający wizualizację podstawowych danych i tworzenie podstawowych statystyk;
- eksportowanie bazy odpowiedzi do plików XLS i CSV – pozwala na dalszą ich obróbkę za pomocą innych narzędzi typu Excel albo pakiety statystyczne.

---

## LimeSurvey

Lime to bezpłatne internetowe narzędzie zaprojektowane specjalnie do tworzenia ankiet internetowych, dostępne na zasadzie *open source*. Wymaga zainstalowania na serwerze – na tym etapie niezbędna jest wiedza informatyczna. Jednak później obsługa samej aplikacji nie wymaga znajomości programowania.

Podstawowe opcje:

- wersja polska wraz z instrukcją;
- tworzenie kont użytkowników i wspólna praca nad ankietami;
- możliwość korzystania z szablonów ankiet;
- wiele różnych typów pytań i formatów odpowiedzi;
- możliwość tworzenia filtrów (tzn. pewne pytania wyświetlają się tylko w przypadku zaznaczenia określonej odpowiedzi na inne pytanie);
- system gromadzenia w postaci bazy danych listy respondentów oraz zarządzania badaniem – tj. zautomatyzowany mechanizm zapraszania do badania, przypominania o badaniu, podziękowań za uczestnictwo, rejestrowania odmów uczestnictwa etc.;
- możliwość kontroli respondentów poprzez system indywidualnych kodów dostępu do ankiety – zapobiega to sytuacji, w której jedna osoba wypełniłaby ankietę kilka razy, a jednocześnie zapewnia pełną anonimowość;
- graficzne przedstawianie podstawowych wyników;
- eksportowanie bazy odpowiedzi w formatach XLS, CSV, SAV.

Więcej: <http://www.limesurvey.org>

---

## 6.2. Zlecenie badań i współpraca z wykonawcą

Jeśli decydujemy się na zlecenie badań zewnętrznemu wykonawcy – czy to będzie firma, czy organizacja pozarządowa, czy też indywidualna osoba – musimy przede wszystkim precyzyjnie określić, jakie są nasze potrzeby i oczekiwania. Zwłaszcza, jeśli skorzystamy z usług instytucji komercyjnej, powinniśmy się nastawić, że co do zasady otrzymamy dokładnie to, co było uzgodnione, nie więcej. Pamiętajmy też, że firmy zajmujące się badaniami społecznymi z najróżniejszych dziedzin – a z reguły są to badania marketingowe – niekoniecznie znają się na specyfice obszaru, którym się zajmujemy. Trochę inaczej może to wyglądać w przypadku, gdy współpracujemy z inną organizacją specjalizującą się w badaniach – ale to też nie jest



reguła. Dlatego punktem wyjścia do zlecenia interesujących nas badań na zewnątrz powinien być tzw. *brief*, czyli dokument, w którym zapiszemy nasze oczekiwania wobec wykonawcy [patrz ramka].

---

### **Schemat *briefu* dla realizatora badań**

Dane zamawiającego (w tym wyznaczona osoba do kontaktu).

Kontekst badania.

Cele projektu, w którym badanie ma być wykorzystane.

Cele badawcze; nasze założenia, hipotezy.

Metodologia – wykorzystania jakich metod oczekujemy.

Produkt z badania – wyników w jakiej formie oczekujemy (baza danych, prezentacja multimedialna, raport w formie dokumentu tekstowego itp.). Jeśli chcemy otrzymać raport z badania, powinniśmy też sprecyzować, jaka powinna być jego zawartość (np. czy oczekujemy tylko analizy wyników, czy również rekomendacji).

Harmonogram badania.

Warunki finansowe.

---

*Brief* może oczywiście być punktem wyjścia do wspólnego ustalenia szczegółów badania. Jeśli nie mamy doświadczenia w badaniach, jest nawet wskazane, żeby realizator pomógł nam wybrać właściwą metodę i podjąć inne ważne decyzje. Dobrą praktyką jest wspólne wypracowywanie pomysłów przez zleceniodawcę – czyli naszą organizację – i realizatora. My wnosimy bowiem wiedzę na temat badanego obszaru i możemy wyjaśnić wszelkie niejasności merytoryczne, natomiast wykonawca wnosi swoje doświadczenie i wiedzę z zakresu samego przebiegu badań. Tego typu wspólna praca przydaje się zarówno na etapie konstruowania narzędzi – np. układania pytań w ankiecie – jak i podczas analizy zgromadzonego materiału badawczego.

Nie jest to łatwe, jeśli nie mamy wiedzy dotyczącej metodologii badań społecznych, ale powinniśmy zawsze kontrolować warunki realizacji badania – a najlepiej samemu je zdefiniować, tak żeby wykonawca musiał spełnić określone kryteria. Powinniśmy być także ostrożni w doborze narzędzi, bowiem może się zdarzyć, że firma, korzystając ze swoistej asymetrii kompetencji, będzie miała skłonność do proponowania przede wszystkim tych narzędzi, które maksymalizują jej wygodę i rentowność, a które niekoniecznie są adekwatne ze względu na postawione

przez nas problemy badawcze. Czasem na potrzeby tak rozumianych negocjacji warto poprosić o konkurencyjną ofertę innej firmy albo choćby incydentalnie o pomoc kogoś, o kim wiemy, że ma kompetencje do bezstronnej oceny zaproponowanej metodologii i kosztorysu.

W tym kontekście szczególnie ważne jest określenie najważniejszych parametrów badania ilościowego. Są to:

- wielkość i dobór próby [patrz podrozdział 7.1.11. „Próba”, str. 85];
- sposób kontaktu z respondentami – wykorzystana technika badawcza (patrz ramka);
- ile razy ankietę ma próbować skontaktować się z respondentem, jeśli jest on niedostępny;
- stopa zwrotu (*response rate*) – czyli ile osób pierwotnie wybranych do badania ostatecznie wzięło w nim udział;
- stopa odmów – czyli ile osób odmówiło udziału w badaniu;
- metody kontroli ankietów, czy faktycznie docierają do wybranych respondentów.

---

## POPULARNE TECHNIKI BADAWCZE

### Ilościowe

PAPI (*Paper and Pencil Interview*) – tradycyjne badanie ankietowe, w trakcie którego ankietę zaznacza odpowiedzi na papierowym kwestionariuszu.

CAPi (*Computer Assisted Personal Interview*) – badanie, w trakcie którego ankietę również spotyka się z respondentem, jednak zapisuje odpowiedzi na komputerze.

CATi (*Computer Assisted Telephone Interview*) – badanie realizowane przez ankietów telefonicznie.

CAWi (*Computer Assisted Web Interview*) – ankietę internetowa, wypełniana samodzielnie przez respondenta.

Badanie audytoryjne – badanie, w trakcie którego ankietę rozdaje kwestionariusze respondentom zgromadzonym w jednym miejscu (np. uczniom w klasie), po czym zbiera je po wypełnieniu.

### Jakościowe

IDI (*Individual In-Depth Interview*), czyli wywiad pogłębiony – wywiad swobodny prowadzony według przygotowanego wcześniej scenariusza, trwa zwykle 30 do 60 minut.

FGI (*Focus Group Interview*), czyli wywiad grupowy bądź „fokus” – wywiad podobny do IDI, ale prowadzony w grupie, z reguły liczącej ok. 8-10 osób, służący zebraniu i porównaniu różnych punktów widzenia na ten sam temat.

---

Bardzo ważne jest także zaznaczenie, że w raporcie z badania oczekujemy zawarcia części metodologicznej – dzięki temu będziemy mieć zebrane w zwartej formie informacje na temat źródła danych. Może nam się to bardzo przydać, bowiem jeśli będziemy powoływać się na własne badania w działalności rzeczniczej, bardzo szybko możemy spotkać się z pytaniami o to, w jaki sposób uzyskaliśmy określone wyniki.

## 7. Analiza i interpretacja danych

### 7.1. ABC statystyki – słowniczek badacza

Prowadzenie rzecznictwa w oparciu o dane wymaga wiedzy o tym, co oznaczają najważniejsze pojęcia używane w analizie. Wbrew pozorom, na podstawowym poziomie nie jest to wiedza szczególnie złożona i nie wymaga ukończenia kursu statystyki czy ekonometrii. Warto więc poznać język, którym posługują się analitycy: nie tylko po to, aby samemu umieć się nim posługiwać, ale także po to, aby móc krytycznie ocenić to, co mówią inni – włączając przedstawicieli ośrodków badawczych, z którymi współpracujemy przy badaniach.

Z myślą o takim zastosowaniu opracowaliśmy krótki słowniczek pojęć, które powinny być jasne dla każdego zainteresowanego wykorzystaniem danych statystycznych w działaniach społecznych. Ich znajomość pozwoli nie tylko precyzyjniej opisywać analizy wykonywane na potrzeby takich działań, ale także zmniejszy naszą podatność na manipulacje ze strony tych, którzy – świadomie lub nieświadomie – wykonują je nierzetelnie.

Część haseł słowniczka dotyczy pojęć związanych ogólnie z analizą danych – bez względu na to, jakie jest ich źródło. Część odnosi się do specyficznego typu danych, których źródłem są badania sondażowe. Postanowiliśmy zwrócić na nie szczególną uwagę z dwóch powodów. Po pierwsze, sondaże są coraz częściej wykorzystywanym narzędziem w działalności rzeczniczej. Znajomość towarzyszącej im terminologii może więc być szczególnie przydatna dla praktyków. Po drugie, jest to też takie źródło danych, którego interpretacji i wykorzystaniu szczególnie często towarzyszą rozmaite kontrowersje, nadużycia czy nieporozumienia.

Nasz słowniczek nie jest w żadnym razie kompletnym kompendium wiedzy o badaniach. To jedynie subiektywny, oparty na naszych własnych doświadczeniach zbiór kluczowych (najczęściej stosowanych, najczęściej błędnie rozumianych) definicji pojęć związanych z analizą danych. Ogranicza się przede wszystkim do pojęć z zakresu statystyki.

Dla wygody każde z haseł opisaliśmy za pomocą tego samego schematu, na który składa się opisowa definicja określonego pojęcia, przykład jego wykorzystania oraz uwagi / ostrzeżenia dotyczące problemów z interpretacją. Staraliśmy się opisać każde hasło zwięźle, unikając przy tym specjalistycznego żargonu i szczegółów nieistotnych dla zrozumienia istoty sprawy. Dla bardziej zaawansowanych czy też dociekliwych użytkowników, ten styl może jednak wydać się nazbyt upraszczający.

Aby pogłębić poszczególne kwestie, zachęcamy do zajrzenia do takich źródeł, jak słownik pojęć statystycznych GUS<sup>34</sup>, dział metodologiczny serwisu Na Straży Sondazy<sup>35</sup>, Elektroniczny Podręcznik Statystyki<sup>36</sup>, a także (dla znających język angielski) filmy Khan Academy<sup>37</sup>. Niektóre pojęcia statystyczne są również dość szczegółowo opisane w Wikipedii.

### 7.1.1. Agregacja

Definicja	Rodzaj operacji matematycznej polegającej na obliczaniu łącznej wartości (np. sumy) określonej cechy w obrębie pewnych grup (np. osób lub podmiotów o podobnych cechach), które to grupy następnie stają się dla nas nową jednostką analizy (to dla nich będziemy prowadzić dalsze obliczenia). Termin, jak wiele w statystyce, powstał przez niezbyt szczęśliwe spolszczenie angielskiego terminu aggregation oznaczającego „skupisko”, „skupienie”.
Przykład	Agregacja danych o powierzchni wszystkich gmin w obrębie poszczególnych województw polega na zsumowaniu powierzchni gmin, które przynależą do określonego województwa. W wyniku tego działania uzyskujemy zagregowany zbiór danych o powierzchni województw.
Warto pamiętać	Najczęściej agregacja polega na sumowaniu wartości, ale może też oznaczać np. uśrednianie ich <b>[patrz podrozdział 7.1.12. „Średnia”, str. 86]</b> lub zastosowanie jakiejś innej funkcji matematycznej.

### 7.1.2. Błąd statystyczny

Definicja	Wskaźnik określający dokładność pomiaru danej cechy oferowaną przez badanie sondażowe – takie, w którym próbujemy opisać interesującą nas zbiorowość czy też populację <b>[patrz podrozdział 7.1.10. „Populacja”, str. 85]</b> , np. społeczeństwo polskie, na podstawie reprezentacji tej zbiorowości, zwanej próbą <b>[patrz podrozdział 7.1.11. „Próba”, str. 85]</b> .  Inaczej mówiąc, jest to błąd pomiaru związany z tym, że próbujemy oszacować częstość lub wielkość jakiejś cechy w całej zbiorowości na podstawie badania jej wybranych przedstawicieli. Mówi o tym, w jakim przedziale najprawdopodobniej mieści się interesujący nas wynik w populacji, którą opisujemy na podstawie zbadanej w sondażu próbki. W gruncie rzeczy jest to oparty o reguły statystyki wariant sytuacji, w której, bazując na naszej wiedzy, „strzelamy”, że wartość danej cechy w całej populacji zawiera się prawdopodobnie w przedziale między wartością X a wartością Y.
-----------	--

34 [http://www.stat.gov.pl/gus/definicje\\_PLK\\_HTML.htm?id=LST-POJ2.htm](http://www.stat.gov.pl/gus/definicje_PLK_HTML.htm?id=LST-POJ2.htm)

35 <http://nastrazysondazy.uw.edu.pl/metodologia>

36 <http://www.statsoft.pl/textbook/stathome.html>

37 <https://www.khanacademy.org/math/probability>

Definicja	<p>Wielkość błędu zależy głównie od liczby respondentów badania (im więcej, tym będzie mniejszy). Liczy się także to, jaki jest rozkład badanej przez nas cechy (na przykład, jak często zdarza się, że respondenci wyrażają zaufanie do innych), to, jak duża część respondentów odpowiedziała na nasze pytanie, a także to, jak wielkie ryzyko nietrafnego przewidywania wyniku jesteśmy w stanie zaakceptować. Zakładając, że ryzyko to nie przekracza 5% i gdyby udało się nam uzyskać dane o wszystkich badanych przez nas podmiotach, błąd ten wynosi nie więcej niż ok. +/- 4,4% dla próby 500-elementowej, ok. +/- 3,1% dla próby 1000-elementowej, ok. +/- 2,2% dla próby 2000-elementowej i ok. +/- 1% dla próby 10 000-elementowej.</p>
Przykład	<p>W badaniu X, przeprowadzonym na losowej, reprezentatywnej próbie 1000 dorosłych mieszkańców Polski 50% respondentów stwierdziło, że ma wpływ na sprawy swojego miasta lub gminy. Biorąc pod uwagę reprezentatywny charakter badania, można powiedzieć, że z 95-procentowym prawdopodobieństwem pogląd ten wyraziło nie więcej niż 53,1% i nie mniej niż 46,9% ogółu mieszkańców Polski.</p>
Warto pamiętać	<p>O błędzie statystycznym możemy mówić tylko wtedy, kiedy mamy do czynienia z badaniem reprezentatywnym na losowej próbie.</p> <p>Każde takie badanie jest obciążone błędem statystycznym.</p> <p>Błąd statystyczny to tylko jeden z czynników mogących wpływać na jakość wyników badania (oprócz niego, istotny jest jeszcze tzw. błąd systematyczny, związany ze sposobem doboru respondentów).</p> <p>Wielkość błędu dla „frakcji” to jednak coś innego niż wielkość błędu dla „średniej” <b>[patrz podrozdział 7.1.12. „Średnia”, str. 86]</b>. Średnia wartość danej cechy również jest obciążona błędem statystycznym, a jego wielkość zależy od skali rozproszenia wartości tej cechy w badanej próbie.</p>
Warto pamiętać	<p>Błąd statystyczny przywołuje się często jako wskaźnik dokładności badania jako takiego. To uproszczenie, w istocie bowiem wielkość błędu nie jest stała – zmienia się w zależności od tego, jak rozkładają się w badanej próbie wyniki dla konkretnego pytania czy konkretnej cechy, której wielkość w populacji chcemy oszacować. Przyjęło się jednak podawać maksymalną wielkość błędu statystycznego „dla frakcji” (frakcja to odsetek badanej próby o określonym poziomie danej cechy, np. odsetek osób ufających innym) jako wartość określającą dokładność badania jako takiego. Jeśli więc czytamy gdzieś, że „błąd statystyczny w badaniu... wyniósł nie więcej niż +/-3%”, oznacza to, że hipotetycznemu wynikowi 50% wśród wszystkich respondentów tego badania odpowiada w całej populacji wynik między 47% a 53%. Dla frakcji mniejszych lub większych niż 50% wielkość błędu się zmniejsza (np. gdyby w przywołanym wyżej przykładowym badaniu zaufanie do innych wyrażało nie 50%, ale 20% ankietowanych, błąd statystyczny dla tego wyniku wynosiłby nie +/- 3,1%, ale +/- 2,5%).</p>

### 7.1.3. Indeks

Definicja	Pojęcie niejednoznaczne, na gruncie badań może oznaczać z jednej strony uporządkowaną listę informacji o elementach bazy danych (to znaczenie niżej pomijamy) z drugiej – syntetyczny wskaźnik, który ma wartość liczbową i wyliczany jest na podstawie kilku innych (tzw. wskaźników cząstkowych) bądź dla każdego elementu zbioru (na przykład dla każdego respondenta badania), bądź dla całego zbioru (jako agregat – na przykład indeks giełdowy).
Przykład	Wskaźnik optymizmu konsumentów: Indeks opracowany przez IPSOS, często wykorzystywany w analizach dotyczących koniunktury gospodarczej w Polsce. Powstaje na bazie jedenastu standardowych pytań zamkniętych zadawanych w comiesięcznym sondażu na próbie 1000 osób. Respondenci oceniają sytuację kraju, swoją sytuację materialną, poziom bezrobocia, inflację, skłonność do zakupu trwałych dóbr konsumpcyjnych i poziom oszczędności w ciągu ostatnich 12 miesięcy oraz dokonują prognozy tych zjawisk na rok następny. Na podstawie tych deklaracji oblicza się wskaźnik optymizmu konsumentów WOK, który może przyjmować wartości od 0 do 200 punktów.
Warto pamiętać	Za każdym wiarygodnym indeksem kryje się zwykle złożona teoria oraz starannie wykalibrowany i sprawdzony model statystyczny. Nie jest więc łatwo stworzyć własny indeks – wymaga to długotrwałej pracy i wielu testów.  Na wartość indeksu składa się zwykle przynajmniej kilka wskaźników cząstkowych. Niekiedy mogą one być dla nas bardziej interesujące niż sam indeks.

### 7.1.4. Istotność statystyczna

Definicja	<p>W raportach z badań często słyszymy, że jakiś wynik jest „istotny statystycznie”. Brzmi to dziwnie – przecież wiele innych wyników omówionych w tych raportach wydaje się nam nie mniej istotnych! Do tego, stwierdzeniom takim często towarzyszą dziwaczne zapisy numeryczne, takie jak np. „<math>p &lt; 0,01</math>”. O co tu chodzi?</p> <p>Istotność statystyczna to termin pojawiający się, kiedy porównuje się uczestników badań sondażowych pod względem natężenia czy występowania pewnych cech (przynajmniej dwóch). Jest to oszacowanie prawdopodobieństwa, z jakim można przyjąć, że różnice w poziomie tych cech nie występują w populacji, z której wylosowano uczestników badania.</p> <p>Po co takie oszacowanie? Jeśli dochodzimy do wniosku, że natężenie jakiejś cechy wśród uczestników badania jest związane z natężeniem innej cechy (na przykład przynależnością uczestników do określonych grup), stajemy przed pytaniem, czy związek ten (nazywany często „zależnością”) charakteryzuje całą populację, z której losowaliśmy uczestników badania, czy też można uznać go za przypadkowy, występujący tylko w naszej próbie? Oczywiście, im wyraźniejsza jest obserwowana zależność, tym mniejsze jest prawdopodobieństwo, że ma ona charakter losowy i nie występuje w populacji.</p>
-----------	---

Definicja	<p>Poziom istotności statystycznej to nic innego jak obliczony na podstawie wyników badania poziom tego prawdopodobieństwa – współczynnik, który oznacza się zwykle literą <math>p</math>. Choć wyraża się go w ułamkach, jest to ni mniej ni więcej, tylko sprowadzony do postaci numerycznej procent szans na to, że w populacji nasz związek nie występuje. Inaczej mówiąc, to prawdopodobieństwo, że nasz wynik nie jest ważny w populacji, którą badamy. Wartość <math>p</math> dąży do 1 (100%) wraz ze wzrostem szans, że mamy do czynienia z przypadkiem lub do 0 (0%) wraz z jego maleniem (patrz przykład). Choć metody wyznaczania tego prawdopodobieństwa różnią się w zależności od tego, jakie cechy porównujemy i jakich narzędzi statystycznych (tzw. testów) używamy, zawsze oznacza ono to samo: szanse na to, że odkryta przez nas zależność ma charakter losowy i nie występuje w populacji, o której wnioskujemy na podstawie naszej próby.</p>
Przykład	<p>W badaniu sondażowym na reprezentatywnej próbie 1000 dorosłych Polaków stwierdzono, że istnieje <b>statystycznie istotny</b> związek [patrz podrozdział 7.1.5. „Korelacja”, str. 81] pomiędzy poziomem zamożności respondentów a ich skłonnością do hazardu, mierzoną liczbą wykupywanych w danym miesiącu zakładów loterii – osoby mniej zamożne częściej grają w totolotka niż osoby zamożne. Poziom istotności dla tego związku wynosi <math>p = 0,03</math>. Oznacza to tylko 3% szans na to, iż nie występuje on w całym polskim społeczeństwie. Inaczej mówiąc, mamy 97-procentową pewność, że jest to związek opisujący nie tylko naszą próbkę, ale całą badaną populację.</p>
Warto pamiętać	<p>O istotności statystycznej możemy mówić w odniesieniu do badań sondażowych na losowych, reprezentatywnych próbach. Tylko losowość doboru respondentów pozwala bowiem przyjąć reguły wnioskowania statystycznego stojące za wartością <math>p</math>.</p> <p>Nie istnieje coś takiego jak zerowy poziom <math>p</math> – po prostu dlatego, że w badaniu opartym na losowej próbie nigdy nie mamy absolutnej pewności, że odkryta przez nas prawdziwość nie jest po prostu dziełem przypadku. Prawdopodobieństwo takie może być jednak niezwykle małe – na przykład nie przekraczać wartości <math>0,0000001</math>, co oznacza <math>0,00001\%</math> szans na błędne uznanie, że występuje w populacji.</p> <p>Istotność statystyczna ma zawsze określoną wartość (prawdopodobieństwo jest z natury określone). Jednak w raportach czytamy zwykle, że albo coś „jest”, albo „nie jest” statystycznie istotne. Powstaje pytanie: od kiedy dany wynik „nie jest istotny”? Jaka musi być wartość <math>p</math>, aby go za taki uznać? W badaniach statystycznych przyjęło się zakładać pewne progi, poniżej których dany wynik traktowany jest jako istotny statystycznie – jednak jest to jedynie rodzaj umowy między badaczami. W sondażach opinii publicznej, takim progiem jest często <math>p=0,05</math>. Jeśli wyliczona istotność statystyczna wyniku testu jest niższa niż ten próg (czyli szansa na to, że wynik ten jest przypadkowy i nie występuje w badanej populacji nie przekracza 5%), uznaje się go za statystycznie istotny. Jeśli przekracza ten próg (nawet o ułamek procenta), większość badaczy nie odważy się traktować go jako istotnego.</p> <p>Poziom istotności statystycznej nie mówi nam o sile opisywanej zależności, ale o prawdopodobieństwie, że istotnie zachodzi ona w całej interesującej nas populacji. Inaczej mówiąc, coś, co jest istotne statystycznie, niekoniecznie jest istotne z punktu widzenia</p>



Warto pamiętać	interpretacji wyników. Z drugiej strony, nawet najbardziej wyraziste zależności obserwowane w naszej próbie mogą być nieistotne statystycznie i wówczas nie powinny nas zajmować. W praktyce poziom istotności statystycznej mocno zależy od liczebności próby w naszym badaniu. Przy bardzo dużych próbach, nawet zupełnie marginalne różnicowania czy związki między respondentami okazują się istotne statystycznie, za to przy niewielkich próbach nawet te z nich, które uznalibyśmy za wstrząsające okazują się nie przekraczać progu istotności statystycznej.
----------------	---

### 7.1.5. Korelacja

Definicja	<p>Termin często używany przez osoby stykające się z wynikami badań, odnoszący się do relacji między wielkościami dwóch lub więcej cech. Korelacja to miara mówiąca o tym, w jakim stopniu zmianom w poziomie jednej z tych cech towarzyszą zmiany w poziomie innej cechy.</p> <p>Korelacja może być dodatnia, kiedy wzrost wartości jednej cechy łączy się ze wzrostem wartości drugiej cechy, lub ujemna, kiedy wzrostowi wartości jednej cechy odpowiada spadek wartości drugiej cechy. Brak korelacji oznacza, że dane dwie cechy są od siebie całkowicie niezależne (niepowiązane ze sobą).</p>
Przykład	W badaniu losów absolwentów uniwersytetu X stwierdzono wyraźną pozytywną korelację między średnią ocen uzyskiwanych przez studentów na studiach a poziomem ich zarobków 5 lat po ukończeniu studiów. Inaczej mówiąc, wyższa średnia na studiach oznacza, statystycznie rzecz biorąc, wyższe zarobki po studiach.
Warto pamiętać	<p>Korelacje to tylko współczynniki mówiące o zależnościach między liczbami. Musimy uważać, aby ich nie nadinterpretować. Do najczęstszych błędów należy przypisywanie danej zależności cech relacji przyczynowo-skutkowej: na przykład uznawanie, że wyższa średnia ocen na studiach sprawia, że zarabiamy więcej po studiach. Tymczasem oba te parametry jedynie współwystępują ze sobą i oba mogą zależeć od innych kwestii (na przykład od tego, jaki był poziom wykształcenia naszych rodziców). Sytuacja, kiedy współwystępowanie zjawisk wynika nie z ich wzajemnej relacji, ale z oddziaływania jakiegoś innego, ukrytego czynnika, nazywana bywa „korelacją pozorną”.</p> <p>Najczęściej wykorzystywaną miarą korelacji jest tzw. współczynnik korelacji liniowej Pearsona. Wiele analiz odwołuje się do niego w sposób błędny, dlatego musimy pamiętać o różnych ograniczeniach jego stosowania. Po pierwsze, jak sama nazwa wskazuje, mierzy on <i>liniową</i> zbieżność natężenia dwóch zjawisk, a więc zakłada stałą (liniową) relację między zmianami wartości jednej zmiennej a zmianami wartości drugiej zmiennej. Może więc prowadzić do błędnych wniosków w sytuacji, kiedy dane zjawiska są ze sobą skorelowane nieliniowo (np. wzrost wartości jednej zmiennej powoduje wzrost wartości drugiej zmiennej, ale tylko do pewnego poziomu). Po drugie, może być stosowany tylko do analizy związków między cechami ilościowymi [<b>patrz podrozdział 7.1.14. „Zmienne ilościowe”, str. 87</b>].</p>

## 7.1.6. Mediana (kwartyle, percentyle)

Definicja	<p>Jedna z miar statystycznych służących do określenia „typowej” czy „przeciętnej” wartości danej cechy w jakimś zbiorze (na przykład w jakiejś grupie). Ścisłej – jest to „średkowa” wartość dla tego zbioru. Można o niej myśleć jako o alternatywie dla znanej wszystkim średniej [patrz podrozdział 7.1.12., „Średnia”, str. 86]. Jeśli wyobrazimy sobie, że szeregujemy wszystkie elementy naszego zbioru ze względu na wartość danej cechy, to mediana jest ni mniej ni więcej, tylko wartością charakteryzującą środkowy element zbioru w przypadku nieparzystej liczby elementów. W przypadku parzystej liczby elementów, jest to średnia arytmetyczna wartości dwóch środkowych elementów. Dlatego często mówi się nie tyle o medianowej wartości cechy, co o wartości cechy dla medianowego elementu (na przykład zarobkach medianowego pracownika).</p> <p>Mediana jest bardzo użyteczną statystyką, ponieważ – w odróżnieniu np. od średniej – jest niewrażliwa na specyfikę rozkładu wartości danej zmiennej. W szczególności nie podlega wpływom związanym z występowaniem w zbiorze skrajnie wysokich wartości. Dlatego czasami (np. w przypadku analiz dochodowych) jest lepszą miarą wartości przeciętnej niż średnia.</p>															
Definicja	<p>Mediana jest jednym z tzw. centyli (czy też percentyli). Na podobnej zasadzie można bowiem wyróżnić miary dla innych niż „środkowy” punktów uszeregowania elementów zbioru. Mówimy więc o kwartylach, czyli miarach wartości cechy dla kolejnych ćwiartek zbioru (pierwszy kwartył to wartość cechy dla elementu wyznaczającego 25% liczebności zbioru, kwartył trzeci to wartość dla elementu wyznaczającego 75% liczebności zbioru. Drugi kwartył to to samo, co mediana), decylach (wartości cechy dla elementów wyznaczających kolejne 10% liczebności zbioru), czy właśnie o centylach (miara znana wszystkim rodzicom).</p>															
Przykład	<p>Jak wynika z badań Stowarzyszenia Klon/Jawor, mediana rocznych przychodów organizacji pozarządowych w Polsce w r. 2012 wyniosła 18 tysięcy złotych. Oznacza to, że medianowa („średkowa”) organizacja pozarządowa osiągnęła właśnie taki przychód. Bazując na tym, można powiedzieć, że przychody połowy organizacji pozarządowych w Polsce nie przekroczyły w tym okresie 18 tys. zł.</p>															
Warto pamiętać	<p>Mediana nie jest wrażliwa na cechy rozkładu wartości danej zmiennej, co ma swoje dobre i złe strony. Z jednej strony, pozwala lepiej przewidywać typową wartość cechy w sytuacji dużych dysproporcji między większością elementów zbioru a kilkoma elementami wyjątkowymi. Z drugiej, może być myląca wtedy, kiedy rozkład cechy jest bardziej płaski. Aby to zrozumieć, rozważmy przykład 10 osób palących papierosy. Poniższa tabela przedstawia 4 różne rozkłady liczby papierosów wypalanych przez te osoby dziennie:</p> <table border="1" data-bbox="296 1188 1042 1279"> <thead> <tr> <th></th> <th>rozkład 1</th> <th>rozkład 2</th> <th>rozkład 3</th> <th>rozkład 4</th> </tr> </thead> <tbody> <tr> <td>Osoba 1</td> <td>5</td> <td>1</td> <td>1</td> <td>1</td> </tr> <tr> <td>Osoba 2</td> <td>5</td> <td>2</td> <td>2</td> <td>2</td> </tr> </tbody> </table>		rozkład 1	rozkład 2	rozkład 3	rozkład 4	Osoba 1	5	1	1	1	Osoba 2	5	2	2	2
	rozkład 1	rozkład 2	rozkład 3	rozkład 4												
Osoba 1	5	1	1	1												
Osoba 2	5	2	2	2												

Warto  
pamiętać

Osoba 3	5	3	3	3
Osoba 4	5	4	4	4
Osoba 5	5	5	5	5
Osoba 6	6	6	6	6
Osoba 7	7	6	7	7
Osoba 8	8	6	8	8
Osoba 9	9	6	9	9
Osoba 10	10	6	10	100

A oto porównanie zachowania mediany i średniej arytmetycznej dla tych rozkładów:

	rozkład 1	rozkład 2	rozkład 3	rozkład 4
Mediana	5,5	5,5	5,5	5,5
Średnia	6,5	4,5	5,5	14,5

Zwróćmy uwagę na kilka ciekawych prawidłowości:

W rozkładzie 3 mediana i średnia są równe – to typowe dla rozkładów idealnie symetrycznych (takich, które koncentrują się równomiernie wokół wartości środkowej).

W rozkładzie 4 mediana spełnia swoją rolę wyraźnie lepiej niż średnia – to dlatego, że liczba papierosów wypalanych dziennie przez osobę 10 jest znacznie większa niż u wszystkich innych osób. Ta jedna osoba zdecydowanie zawyża wartość średniej.

W rozkładach 1 i 2 nieznacznie lepiej zachowuje się średnia. Mediana pokazuje, że połowa badanych wypala nie więcej niż 5,5 papierosa dziennie. Nie zdaje jednak sprawy z tego, że połowa badanej grupy pali albo tylko o pół papierosa mniej (1 rozkład) albo tylko o pół więcej (2 rozkład). Średnia jest pod tym względem bardziej czuła.

### 7.1.7. Modalna (Dominanta)

Definicja	Najczęściej występująca w danym zbiorze wartość określonej cechy
Przykład	W zbiorze liczb {1,2,2,3,4,5,6} modalną jest 2.
Warto pamiętać	Rozkład może mieć więcej niż jedną wartość modalną. Wartość „najczęstsza” nie musi być wcale częsta. Warto o tym pamiętać zwłaszcza, gdy analizujemy zjawiska o charakterze ilościowym i o dużym zakresie wartości. Np. informacja o najczęstszej wielkości zadłużenia jednostek samorządu terytorialnego będzie interesująca, jeśli wartości tej zmiennej są jakoś pogrupowane, ale nie wtedy, kiedy wyrażone są np. postaci dokładnej wielkości zadłużenia w złotych lub procentach.

### 7.1.8. Odchylenie standardowe, wariancja

Definicja	<p>Odchylenie standardowe to miara rozproszenia wyników w danym zbiorze (np. w jakiejś grupie) wokół średniej <b>[patrz podrozdział 7.1.12. „Średnia”, str. 86]</b>. Jego wartość uzyskuje się, obliczając średnią (kwadratową) z różnic pomiędzy wielkością jakiejś cechy dla każdego z elementów zbioru a średnią arytmetyczną tej cechy w całym zbiorze. Im mniejsze odchylenie standardowe, tym „bliżej siebie” są poszczególne elementy zbioru pod względem danej cechy (tym mniejsza jest przeciętna odległość ich wyników od wartości średniej).</p> <p>Wartość odchylenia standardowego podniesiona do kwadratu to wariancja.</p> <p>Odchylenie standardowe (podobnie jak wariancja) może przyjmować wyniki od 0 do plus nieskończoności, gdzie 0 oznacza całkowity brak różnic między wszystkimi elementami zbioru.</p>
Przykład	<p>W klasie liczącej 10 uczniów przeprowadzono 2 egzaminy. Oceny z pierwszego egzaminu to {4,4,4,4,4,4,4,4,4,5}, a z drugiego {2,3,3,3,4,4,4,6,6,6}. Choć średnia arytmetyczna ocen z obu egzaminów jest taka sama (4,1), to widać, że w drugim egzaminie oceny były bardziej zróżnicowane. Potwierdza to wielkość odchylenia standardowego, które dla pierwszego egzaminu wynosi 0,3 stopnia, a dla egzaminu drugiego 1,4 stopnia.</p>
Warto pamiętać	<p>Obliczanie odchylenia standardowego ma sens w przypadku zmiennych ilościowych <b>[patrz podrozdział 7.1.14. „Zmienne ilościowe”, str. 87]</b>.</p> <p>Wielkość odchylenia standardowego wyraża się w jednostkach zmiennej, dla której to odchylenie wyliczono (w powyższym przykładzie – w stopniach).</p>

### 7.1.9. Operat losowania

Definicja	<p>Pojęcie często spotykane w opisach badań sondażowych. Oznacza zbiór (wykaz, listę) zawierający informacje o wszystkich elementach danej populacji <b>[patrz podrozdział 7.1.10. „Populacja”, str. 85]</b>, z którego dobiera się (losuje) określoną liczbę elementów do badania.</p>
Przykład	<p>Operatem badań sondażowych opinii publicznej w Polsce (wykonywanych na losowej próbie adresowej) jest zwykle rejestr PESEL (Powszechny Elektroniczny System Ewidencji Ludności).</p>
Warto pamiętać	<p>Operat to inaczej spis wszystkich jednostek, które powinny mieć równe szanse na włączenie do próby badawczej <b>[patrz podrozdział 7.1.11. „Próba”, str. 85]</b>. Każda taka jednostka może występować w operacie tylko raz (inaczej prawdopodobieństwo jej wylosowania byłoby większe niż dla innych jednostek). Optymalnie, operat powinien zawierać dane o wszystkich jednostkach tworzących badaną zbiorowość (populację).</p>

### 7.1.10. Populacja

Definicja	Zbiorowość, której dotyczy dane badanie – ta, do której mają odnosić się jego wyniki.
Przykład	Populacją, do której odnosi się większość badań opinii publicznej w Polsce, jest zbiorowość wszystkich mieszkańców Polski powyżej 15. roku życia.
Warto pamiętać	W potocznym języku słowo „populacja” bywa z reguły używane w odniesieniu do zbiorowości ludzkich lub zwierzęcych. W badaniach odnosi się ono do dowolnych elementów stanowiących zbiór, którego dotyczą wyniki badań. Możemy więc mieć do czynienia z populacją samorządów terytorialnych, firm rodzinnych lub kół zębatach (jeśli planujemy badanie próbkowe mierzące do kontroli jakości tych kół).

### 7.1.11. Próba

Definicja	<p>Zbiorowość uczestników danego badania. Próbę wybiera się (najczęściej z zastosowaniem procedur mających zapewnić losowość wyboru) w celu uzyskania określonej wiedzy nie o samej próbie, ale o populacji, z której ona pochodzi. Przy zastosowaniu odpowiednich procedur doboru próby i realizacji badania, możemy w ten sposób uzyskać wiedzę o populacji bez konieczności badania wszystkich jej przedstawicieli.</p> <p>W badaniach społecznych wyróżnia się różne rodzaje prób. Jeśli zbadamy wszystkich członków danej populacji (czasami taka strategia jest uzasadniona), mówimy o próbie pełnej. Jeśli „ręcznie” wybieramy jej przedstawicieli na podstawie jakichś kryteriów, mamy do czynienia z próbą celową. Próba, której struktura ma odzwierciedlać strukturę populacji (np. liczebność próby pod względem płci i wieku ma być podobna jak liczebność populacji), nazywana jest próbą reprezentatywną. Jeśli do tego skład próby ustalany jest w sposób losowy, mamy do czynienia z próbą losową lub losową próbą reprezentatywną.</p>
Przykład	<p>Komunikat „Badanie przeprowadzono na reprezentatywnej, losowej próbie 1000 mieszkańców Polski w wieku powyżej 15 lat” oznacza, że uczestnicy badania zostali dobrani w sposób losowy (za pomocą matematycznych technik zapewniających wszystkim jednostkom w operacji badawczej ustalone prawdopodobieństwo bycia wybranymi do badania) z operatu, na który składali się wszyscy mieszkańcy Polski powyżej 15. roku życia. W badaniu uczestniczyło 1000 osób, a struktura tej grupy odpowiada ze względu na istotne dla badania kryteria (np. płeć, wiek, miejsce zamieszkania) strukturze populacji wszystkich osób powyżej 15. roku życia w Polsce.</p> <p>Badanie na próbie zawsze wiąże się z koniecznością szacowania wyników badania dla całej populacji. Wyniki takie są więc zawsze obarczone pewnym błędem <b>[patrz podrozdział 7.1.2. „Błąd statystyczny”, str. 77]</b>.</p> <p>Wielkość próby podawana w komunikatach badawczych to z reguły tzw. wielkość „próba zrealizowana”, czyli liczba podmiotów (np. osób), które rzeczywiście wzięły udział w badaniu. W praktyce liczba ta stanowi jedynie część (zwykle mniejszą) liczby pod-</p>

Warto pamiętać	<p>miotów zaproszonych do badania, którzy z różnych powodów ostatecznie w nim nie uczestniczą. Im mniejsza jest liczebność próby zrealizowanej w stosunku do wylosowanej, tym mniej wiarygodne stają się wyniki badania.</p> <p>Próba reprezentatywna nie musi mieć charakteru losowego – to często łączone, ale rozłączne pojęcia. Przykładem nielosowej próby reprezentatywnej może być tzw. próba kwotowa, polegająca na zbadaniu odpowiedniej liczby respondentów należących do określonych grup czy kategorii, w proporcjach określanych w stosunku do całej populacji, której dotyczy badanie (np. odpowiadającej udziałowi w populacji liczby osób o określonym wieku i określonej płci). Taka próba spełnia kryteria reprezentatywności, ale nie jest losowa, ponieważ o doborze respondentów nie decyduje wyłącznie ślepy traf (mają na niego wpływ np. ankieterzy, dostępność respondentów w danym miejscu lub czasie itp.). Nielosowy charakter próby sprawia, że nie możemy precyzyjnie oszacować dokładności wyników badania, nawet jeśli jest reprezentatywna. Takie wyniki trudniej jest traktować jako obraz sytuacji w populacji, z której pochodzi próba.</p>
----------------	---

### 7.1.12. Średnia (arytmetyczna)

Definicja	Jedna z najpopularniejszych miar służących ustaleniu przeciętnej wartości danej cechy w określonej grupie czy zbiorze. Oblicza się ją, sumując wartość tej cechy dla wszystkich elementów zbioru, a następnie dzieląc przez jego liczebność.
Przykład	Według szacunków GUS przeciętne (średnie) wynagrodzenie w gospodarce narodowej w 2012 roku wyniosło 3521,67 zł. Wartość tę uzyskano, sumując (na podstawie badań przedsiębiorstw) wszystkie uzyskane w danym roku wynagrodzenia osobowe (bez osób wykonujących pracę nakładczą i zatrudnionych za granicą), honoraria, wypłaty z tytułu udziału w zysku i dodatkowe wynagrodzenia roczne w przypadku pracowników sfery budżetowej, a następnie dzieląc tę sumę przez przeciętną liczbę osób zatrudnionych w roku.
Warto pamiętać	<p>Średnia arytmetyczna jest bardzo wrażliwa na wartości odstające i skrajne, czyli sytuacje, kiedy wyniki niewielkiej części danego zbioru znacznie różnią się od wyników pozostałych elementów <b>[patrz przykład dotyczący mediany › podrozdział 7.1.6., str. 82]</b>. Jest to typowe dla wielu zjawisk (na przykład rozkładów dotyczących przychodów), co – jak widać po powyższym przykładzie – nie zawsze jest uwzględniane w badaniach. Aby ograniczyć wpływ takich sytuacji na oszacowania, stosuje się rozmaite zabiegi – na przykład zastępuje średnią arytmetyczną medianą <b>[patrz podrozdział 7.1.6. „Mediana”, str. 86]</b> albo oblicza się tzw. średnią obciążoną – z wyłączeniem przypadków o najwyższych i najniższych wartościach danej cechy.</p> <p>Choć pojęcie „średnia” zwykle odnosi się do średniej arytmetycznej, nie jest to jedyny rodzaj średniej. Np. w badaniach szeregów czasowych, kiedy analizuje się przeciętne tempo zmian jakiegoś zjawiska, powszechnie stosowana jest tzw. średnia geometryczna (stosujemy ją, kiedy chcemy wyrównać zakresy zmiennych z różnych pomiarów, dla których liczymy średnią). W niektórych innych zastosowaniach (np. przy obliczaniu średniej dla wartości stosunkowych – takich jak prędkość w km/h) wykorzystuje się średnią harmoniczną.</p>

### 7.1.13. Ważenie

Definicja	Pojęcie odnoszące się do sytuacji, w której chcemy, aby nasze obliczenia uwzględniały różny „wkład” poszczególnych elementów naszego zbioru w badane zjawisko. Na przykład w badaniach opinii publicznej zwykle struktura próby badawczej <b>[patrz podrozdział 7.1.11. „Próba”, str. 85]</b> w mniejszym lub większym stopniu odbiega od struktury populacji, do której badanie się odnosi. Aby wziąć to pod uwagę, stosuje się „wagi”, czyli mnoży liczebność poszczególnych elementów zbioru (osób) przez współczynnik, który koryguje te różnice (taki współczynnik to właśnie „waga”).
Przykład	Wyobraźmy sobie badanie dotyczące zarobków, w którym udało nam się zbadać dwukrotnie za mało kobiet w stosunku do populacji, której dotyczy badanie. <b>[patrz podrozdział 7.1.10. „Populacja”, str. 85]</b> .  Zakładając, że zarobki kobiet są niższe niż mężczyzn, można przypuszczać, że przeciętna wartość zarobków uzyskana na podstawie takiego badania będzie zawyżona w stosunku do wartości występującej w populacji (w badaniu uczestniczy więcej mężczyzn niż powinno, a mężczyźni mają przeciętnie wyższe zarobki). W takiej sytuacji, możemy założyć, że każda kobieta uczestnicząca w badaniu będzie w nim liczona podwójnie, a każdy mężczyzna zostanie „podzielony” przez dwa.
Warto pamiętać	Jeśli chcemy zwiększyć znaczenie danego elementu w wynikach, waga będzie większa niż 1, jeśli zmniejszyć – mniejsza niż 1, ale większa niż 0 (pomnożenie przez 0 oznacza w istocie usunięcie danego elementu z analizy – jego znaczenie dla wyników staje się zerowe).

### 7.1.14. Zmienne ilościowe (interwałowe, ilorazowe) i jakościowe (nominalne, porządkowe)

Definicja	Dane, które chcemy wykorzystywać w naszych analizach, różnią się od siebie własnościami. Różnice te dotyczą zawartości tych danych czy też bogactwa informacji, która się w nich kryje. Jeśli na przykład podzielimy jakąś grupę na cztery podgrupy, wiemy tylko o tym, do której z podgrup należą jej przedstawiciele (zmienną określającą taką przynależność nazywamy nominalną – ponieważ „nominuje” ona dany przypadek do określonej kategorii, ale nie mówi nic więcej). Jeśli jednak wiemy także, że grupy te wyodrębnione zostały na podstawie natężenia jakiejś cechy (na przykład zachorowalności na choroby serca), potrafimy uporządkować członków całego naszego zbioru ze względu na tę cechę (stąd zmienną taką nazywamy porządkową). Jeśli do tego wartości tej zmiennej mówią nam o tym, ile zawałów serca przeszedł każdy przedstawiciel grupy (grupa 1 to 0 zawałów, grupa 4 = 3 zawały), możemy nie tylko uporządkować ich ze względu na poziom zachorowalności, ale także stwierdzić, że przedstawiciele każdej grupy różnią się od siebie o taki sam interwał (dlatego takie zmienne nazywa się interwałowymi), a do tego stwierdzić, że natężenie badanej cechy w grupie 4 jest 3 razy większe niż w grupie 1 (takie zmienne nazywa się ilorazowymi).
-----------	---

Definicja	Zmienne nominalne i porządkowe nazywa się inaczej zmiennymi jakościowymi, a interwałowe i ilorazowe – ilościowymi. Im większą precyzję informacji oferuje nam dana zmienna, tym bogatszy jest wachlarz analiz statystycznych, którym możemy ją poddać. Przykładowo, obliczenie średniej dla zmiennej określającej w kategoriach liczbowych przynależność członków grupy do poszczególnych województw (1 = dolnośląskie, 2 = kujawsko-pomorskie itp.) nie ma sensu.
Przykład	Zmienna nominalna – płeć Zmienna porządkowa – wykształcenie Zmienna interwałowa – data urodzenia Zmienna ilorazowa – wiek
Warto pamiętać	W badaniach często spotykamy się ze stosowaniem analiz statystycznych nieadekwatnych dla zmiennych, które są przedmiotem analizy. Typowym przykładem może być stosowanie średniej dla obliczania przeciętnej wartości zmiennych porządkowych (np. obliczanie przeciętnego poziomu zaufania na podstawie pytania o to, czy ktoś bardziej czy mniej ufa ludziom). Warto pamiętać, że wartość takich analiz jest ograniczona.  Zmienne (czy też skale) jakościowe to nie to samo, co badania jakościowe. Podział na zmienne ilościowe i jakościowe dotyczy w gruncie rzeczy tylko informacji zapisanych w postaci numerycznej, a więc charakterystycznych dla badań ilościowych.

### 7.1.15. Zmienne zależne/niezależne

Definicja	Termin używany w sytuacji, kiedy badacze próbują określić związek między jakimiś cechami i zakładają przyczynowo-skutkową relację między nimi. Zmienna zależna nazywana jest inaczej objaśnianą – to ta, której zmiany chcemy wyjaśniać za pomocą informacji o zmianach innych zmiennych. Zmienna niezależna (objaśniająca) to ta, która ma służyć do wyjaśniania zmian w wartości zmiennej zależnej (objaśnianej).
Przykład	Badacze podjęli próbę oszacowania wpływu wykształcenia na sytuację ekonomiczną. Wykształcenie jest w tym przykładzie zmienną wyjaśniającą, a sytuacja ekonomiczna (a ściślej – jej miara, na przykład wartość majątku) zmienną wyjaśnianą.
Warto pamiętać	W badaniach społecznych relacja między zmienną wyjaśnianą i wyjaśniającą jest bardzo często z góry zakładana przez badacza. Czasami taki związek narzuca się w sposób oczywisty i ma charakter naturalny – np. wtedy, kiedy stwierdzamy, że wykształcenie dzieci jest związane z wykształceniem rodziców (tu: zależy od niego). Jednak nie zawsze jest on tak oczywisty. W powyższym przykładzie zakłada się, że majątek zależy od wykształcenia, choć równie dobrze można by przyjąć, że jest odwrotnie: to poziom wykształcenia zależy od zamożności. Warto mieć to na uwadze, czytając wszelkie opracowania statystyczne w dziedzinie nauk społecznych.



## 7.2. Rzetelność danych w rzecznictwie – na co zwracać uwagę?

Najgorszą rzeczą, jaka może przydarzyć się wszystkim, którzy korzystają z danych w działaniach rzeczniczych, jest zawierzenie informacjom, które okazują się nierzetelne czy też zmanipulowane. Ujawnienie mankamentów danych, na których budujemy przekaz rzeczniczy, może oznaczać pogrzebanie sprawy, którą staramy się promować, a w konsekwencji wizerunkową śmierć naszej organizacji. Dlatego poszukując faktów mogących wesprzeć nasze stanowisko, a także próbując zebrać te fakty samodzielnie (prowadząc własne badania), dobrze jest przestrzegać kilku sprawdzonych zasad, które nie tylko pozwolą ustrzec się takich wpadek, ale także pomogą przygotować się na wszelką krytykę ze strony oponentów promowanej przez nas idei.

Część z niżej wymienionych zasad ściśle dotyczy sytuacji, w której samodzielnie analizujemy lub gromadzimy dane. Inne odnoszą się raczej do zasad oceny jakości danych dostarczanych nam przez kogo innego. Wszystkie odnoszą się jednak przede wszystkim do tzw. danych ilościowych, a więc mających charakter statystyczny (choć część może być z powodzeniem zastosowana również w odniesieniu do danych jakościowych). Spis tych zasad jest jedynie pewną propozycją ujęcia tematu – bazuje na naszych własnych doświadczeniach dotyczących różnych sposobów manipulacji danymi lub błędów, na które sami natknęliśmy się w praktyce badawczej. Doświadczenia te podzieliliśmy na siedem zagadnień związanych z procesem pozyskiwania i analizy danych – nazwaliśmy je siedmioma „grzechami głównymi”. Można je traktować jako rodzaj listy kontrolnej, którą warto stosować, gdy chcemy odwołać się w naszej narracji do jakiegokolwiek wyniku badań.

### 7.2.1. Grzech pierwszy: nieobiektywność

Jedyną rzeczą gorszą niż przywoływanie błędnych danych jest wykorzystywanie danych celowo zmanipulowanych lub narażających nas na zarzut manipulacji. Zastrzeżenia dotyczące obiektywności danych są zresztą jednym z podstawowych problemów, z którymi muszą zmierzyć się organizacje chcące promować, w oparciu o fakty, jakieś wyraziste stanowisko. Domena badań i analiz jest kojarzona przede wszystkim z rodzajem akademickiego dystansu, który nie przystaje do logiki działania takich organizacji. Niezależnie więc od tego, jak skrupulatnie i uczciwie będziemy używać danych, nasza argumentacja będzie tym częściej kwestionowana, im większe kontrowersje wywołuje podejmowany przez nas problem. Tym bardziej więc musimy dołożyć starań, by nie korzystać z danych budzących tego rodzaju wątpliwości. Jakie to dane?

- **Wyjęte z kontekstu:** przemawiające na rzecz naszej tezy, ale omawiane lub dostępne publicznie wraz z innymi danymi, które tej tezie przeczą (przykładem może być artykuł Rzeczpospolitej na temat losów 6-latków w związku z obniżeniem wieku obowiązku szkolnego)<sup>38</sup>.
- **W sposób oczywisty zmanipulowane:** na przykład bazujące na tendencyjnie zadawanych, niezgodnych ze sztuką pytaniach sondażowych (przykład z sondażu przeprowadzonego w czerwcu 2013 roku przez Instytut Badania Homo-Homini dla stowarzyszenia Republikanie: „Gdyby w najbliższy weekend odbyły się wybory parlamentarne i stowarzyszenie Republikanie brało w nich udział, na jakie ugrupowanie oddał(a)by Pan(i) swój głos?”)<sup>39</sup>.
- **Sponsorowane:** finansowane przez sponsora mającego wyraźny ideologiczny lub ekonomiczny interes w uzyskaniu określonych wyników – niezależnie od tego, czy fundator rzeczywiście wpływał na wyniki badań (na przykład badania dotyczące szkodliwości palenia papierosów sponsorowane przez producentów tytoniu).

### 7.2.2. Grzech drugi: nieaktualność

Ciosem dla wiarygodności naszego przekazu może też być sytuacja, kiedy odwołujemy się do danych zgromadzonych zbyt dawno – zwłaszcza, jeśli dane zagadnienie podlega dynamicznym przemianom i istnieje ryzyko, że stare badania tych przemian nie uwzględniają. Nieaktualność to zatem względne pojęcie – czasami, w przypadku szczególnie drażliwych tematów, może zostać wytknięta nawet obiektywnie całkiem świeżym danym, ale nie tym najaktualniejszym. Warto się więc upewnić, czy dane, do których się odwołujemy, są najnowszym dostępnym omówieniem interesującego nas zagadnienia.

### 7.2.3. Grzech trzeci: niekompletność, irrelewantność danych

Kolejnym często spotykanym zarzutem wobec kampanii rzeczniczych jest oskarżenie o niekompletność przywoływanych danych. Tę cechę można rozumieć w kategoriach podmiotowych – jako odwoływanie się do danych niedotyczących całej populacji, której określone stanowisko dotyczy, jak również w kategoriach przedmiotowych – kiedy analizy, na które się powołujemy, obejmują tylko wycinek problemu albo nie dotyczą go bezpośrednio. Dość sugestywnym przykładem tego rodzaju zarzutów są losy kampanii Fundacji „Maciuś” dotyczącej niedożywienia wśród dzieci.

38 <http://www.ibe.edu.pl/pl/media-prasa/aktualnosci-prasowe/197-zla-interpretacja-wynikow-analiz-ibe-nie-moze-podwazac-reformy>

39 <http://www.homohomini.com.pl/news/130/55/Badanie-dla-Stowarzyszenia-Republikanie/d,Aktualno%C5%9BCi>

Na początku roku 2013 Fundacja wywołała poruszenie raportem, z którego wynikało, że ponad 800 tysięcy dzieci w Polsce cierpi głód. Dyskusja, którą wywołała ta kampania, szybko zamieniła się w spektakularną nagonkę na organizatorów badania (przy okazji ujawniono inne, niezwiązane z kampanią praktyki organizacji, które skutecznie podważyły jej reputację). Okazało się, że oszacowania dotyczące niedożywienia wśród dzieci oparte są nie na danych dotyczących samych dzieci, ale na opiniach blisko 800 pracowników pomocy społecznej oraz szkół, zebranych w sondażu telefonicznym. Podstawą feralnych oszacowań były deklaracje tych osób dotyczące proporcji dzieci cierpiących z powodu niedożywienia w ich otoczeniu. Oczywiście jest przy tym, iż uczestnicy badania w żaden sposób nie mogli posiadać precyzyjnych danych dotyczących skali niedożywienia. W dodatku ich deklaracje odnosiły się jedynie do dzieci uczęszczających do pierwszych klas szkoły podstawowej, podczas gdy oparte na nich wnioski sugerują, że szacunki te zostały w sposób nieuprawniony (bo nieoparte na żadnych empirycznych podstawach) przeniesione również na dzieci starsze, do 18. roku życia (tylko wówczas bowiem wyniki badania przekładają się na liczbę 800 tysięcy niedożywionych dzieci).

#### 7.2.4. Grzech czwarty: błędy metodologiczne

Nieszczęśliwe rozstrzygnięcia różnych dylematów metodologicznych związanych z procesem gromadzenia danych mogą mieć katastrofalne skutki dla ich jakości. Poniżej wymieniamy najważniejsze rodzaje takich błędów. Niezależnie od tego, czy sami prowadzimy badania, czy też korzystamy z dorobku innych, powinniśmy upewnić się (w takim stopniu, w jakim jest to możliwe), że nasze dane są od nich wolne.

#### **Zły dobór techniki realizacji badań**

Ten bardzo często spotykany błąd polega na takim doborze technik pozyskiwania danych, który jest nieoptymalny z punktu widzenia precyzji bazujących na tych danych oszacowań. Może tak być w sytuacji, kiedy usiłujemy badać problem, odwołując się do technik sondażowych, podczas gdy istnieją dane dobrze opisujące go na poziomie całej populacji. Częściej jednak spotykamy się z nim na gruncie samych badań, kiedy decydujemy się na przeprowadzenie sondażu, korzystając z niedostosowanych do specyfiki danej populacji technik realizacyjnych – na przykład sondażu telefonicznego w sytuacji, kiedy zbieramy dane wrażliwe; sondażu typu CAWI (ankiety online) wtedy, kiedy znaczna część naszej populacji nie ma dostępu do komputera/Internetu itp. O wyborze takich technik decydują najczęściej względy ekonomiczne, jednak nie zawsze usprawiedliwiają one tego rodzaju decyzję – mogą sprawić, że mniejszym kosztem uzyskamy dane, które nie będą miały żadnej rzeczywistej wartości.

### **Nieodpowiednia liczebność próby/błędy doboru próby**

Jeśli mamy do czynienia z badaniami sondażowymi (tzw. badaniami na próbie), upewnijmy się, że liczebność i konstrukcja próby badawczej nie mają wpływu na wyniki. Nagminnym błędem jest wnioskowanie o zmianach w natężeniu jakiegoś zjawiska lub porównania między grupami na podstawie nie dość licznej próby. Regularnie obserwujemy to w przypadku analiz sondaży poparcia dla partii politycznych, które opierają się zwykle na 500-elementowych próbach. W takich sondażach przesunięcia czy dysproporcje na poziomie 2 czy 3 punktów procentowych nie dają żadnych podstaw do dyskusji dotyczących fluktuacji poparcia w populacji, której dotyczy badanie.

Pamiętajmy więc, że jeśli chcemy stawiać takie tezy, różnice, które obserwujemy, muszą być istotne statystycznie [patrz podrozdział 7.1.4. „Istotność statystyczna”, str. 79]. Mówiąc najprościej, muszą być oparte na dość licznej próbie, aby można było uznać je za wiarygodne. Co to znaczy „dość liczna próba”? Nie ma jednej odpowiedzi na to pytanie, chodzi bowiem nie tylko o jej liczebność. W stwierdzeniu, czy mamy do czynienia z godnymi odnotowania zróżnicowaniami, mogą nam jednak pomóc dostępne w Internecie kalkulatory istotności statystycznej (np. na stronie [www.obop.com.pl](http://www.obop.com.pl)). Jeśli nie potrafimy ich użyć, poszukajmy potwierdzenia interesujących nas różnic w raporcie omawiającym dany zbiór danych albo zwróćmy się do specjalisty z prośbą o pomoc w ich interpretacji.

Liczy się nie tylko wielkość próby – nawet te bardzo liczne mogą nie gwarantować uzyskania rzetelnych, precyzyjnych wyników, jeśli ich dobór oparty jest na złych założeniach. Często przywoływanym przykładem takiej sytuacji jest olbrzymie badanie preferencji przedwyborczych przeprowadzone w Ameryce lat 30. przez popularne wówczas czasopismo „Literary Digest”. Choć w przeprowadzonym metodą pocztową badaniu wzięło wówczas udział 2,3 mln respondentów, wyniki były odległe od rzeczywistego rozkładu głosów w wyborach (nietrafnie przewidziano wyraźne zwycięstwo A. Landona nad F. Rooseveltem). Przyczyną okazał się dobór próby w oparciu o książki telefoniczne i rejestr pojazdów, co spowodowało niedostateczną reprezentację w próbie przedstawicieli klasy robotniczej, w znacznej części niemającej wówczas dostępu do takich dóbr. Wbrew pozorom, takie sytuacje zdarzają się również we współczesnej praktyce badawczej – na przykład, gdy w badaniach statystyki publicznej uwzględnia się jedynie podmioty gospodarcze zatrudniające więcej niż 9 osób, w badaniach dotyczących problemów społecznych bierze się pod uwagę tylko osoby będące odbiorcami pomocy społecznej itp.

## **Zmiany metodologii pomiaru wpływające na porównywalność danych w czasie**

Bardzo często jesteśmy zaskakiwani danymi, które wskazują na wyraźny wzrost lub spadek natężenia jakiegoś zjawiska w czasie. Jeśli wcześniej nie podlegało ono tak wyraźnej dynamice, warto upewnić się, że nie są temu winne zmiany w metodologii pomiaru tego zjawiska lub inne istotne dla porównań zmiany. Przykładem tego rodzaju sytuacji mogą być gorące dyskusje wokół publikacji GUS ze stycznia 2012 dotyczącej wzrostu wynagrodzeń<sup>40</sup>. Wskazywała ona na znaczący wzrost poziomu wynagrodzeń na przełomie lat 2011/2012 – tymczasem według analityków i ekspertów źródłem tej dynamiki były przede wszystkim zmiany w sposobie doboru próby w badaniach GUS oraz zmiany podatkowe, które skłoniły część przedsiębiorców do wcześniejszego wypłacania dodatkowych wynagrodzeń/dywidend.

## **Błędy w konstrukcji narzędzi służących zbieraniu danych**

Jednym z najpowszechniejszych czynników zniekształcających wyniki badań są błędy w konstrukcji narzędzi badawczych, które służą do ich zbierania (przede wszystkim zdarza się to w badaniach sondażowych i dotyczy zadawanych pytań). Mogą się one zdarzyć nawet w największych, najbardziej prestiżowych badaniach, dlatego chcąc mieć pewność, że używamy rzetelnych danych (albo w sytuacjach, w których mamy podejrzenie co do danych podawanych publicznie przez innych i chcemy je zakwestionować), warto przyjrzeć się bliżej pytaniom, na których bazujemy. O wyniku badania może decydować wiele elementów budowy kwestionariusza: używane słownictwo (im mniej neutralne, tym bardziej wpływa na percepcję badanych), możliwości odpowiedzi towarzyszące tzw. pytaniom zamkniętym (z gotowymi odpowiedziami do wyboru), kolejność pytań, kolejność możliwych odpowiedzi, a nawet to, czy respondenci są informowani o opcji niedopowiadania w ogóle.

## **Błędy w realizacji badań**

Dotyczą one przede wszystkim badań ankierskich. Nawet najlepiej zaprojektowane badanie tego typu nie będzie wiarygodne, jeśli nie zostanie profesjonalnie przeprowadzone. Tego rodzaju problemy są jednak bardzo trudne do wykrycia – zwykle się o nich nie informuje. Pamiętajmy jednak o nich, prowadząc własne badania lub zlecając je zewnętrznym wykonawcom. Upewnijmy się, że:

- respondenci naszego badania naprawdę pochodzą z populacji, której badanie dotyczy;
- liczba odmów udziału w badaniu nie jest za wysoka (im jest niższa, tym lepiej);

---

<sup>40</sup> <http://biznes.pl/magazyny/finanse/belka-wzrost-plac-umiarkowany-styczniovy-wzrost-ii,5038504,magazyn-detel.html>

- praca związana ze zbieraniem danych (np. przez ankierów) jest odpowiednio kontrolowana przez koordynatorów badania;
- ankierzy zostali odpowiednio przeszkoleni pod kątem realizacji badania, znają jego cel i pracują zgodnie z obowiązującymi wytycznymi (spis standardów i zasad kontroli badań ankierskich można znaleźć na stronie Organizacji Firm Badania Opinii i Rynku [www.ofbor.pl](http://www.ofbor.pl), w zakładkach *Standardy* oraz *PKJPA<sup>41</sup>*).

### 7.2.5. Grzech piąty: błędy techniczne i obliczeniowe

Obok rozstrzygnięć metodologicznych, na jakości wykorzystywanych przez nas danych mogą zaważyć także błędy techniczne. Są to przede wszystkim problemy związane z nieprawidłowym działaniem narzędzi badawczych albo niewłaściwą konstrukcją baz danych zawierających wyniki badań. Ujawniają się one w postaci rozmaitych nielogiczności lub sprzeczności w interpretacji tych wyników. Występowanie takich niespójności powinno wzbudzić naszą czujność. Mogą one oznaczać na przykład problemy związane z niejednoznacznym określeniem sposobu wpisywania przez respondentów lub ankierów danych liczbowych (część badanych może stosować zaokrąglenia, na przykład podając kwoty w tysiącach, część może przywoływać pełne wartości), niejednoznaczną identyfikacją braków danych (np. kiedy braki danych w przypadku informacji o wieku respondenta w latach kodowane są za pomocą kodu 99), dublowanie informacji w zbiorze (wielokrotne występowanie w nim tych samych obserwacji, np. respondentów), niespójności danych wskazujące np. na nieprawidłowe działanie tzw. reguł przejścia w kwestionariuszach albo na inne problemy związane ze spójnością bazy danych (autentyczne przykłady: organizacje pozarządowe nieprowadzące działalności gospodarczej, które twierdzą, że uzyskały z niej przychód; odbiorcy pomocy społecznej, którym w kolejnych odsłonach badania zmieniają się daty urodzenia).

Sprawdźmy też, czy nasze obliczenia nie są oparte na fałszywych podstawach. Przeanalizujemy rozkład wyników, które promujemy, sprawdźmy, czy nie występują w nim niepokojące nieciągłości lub wyniki odstające. Upewnijmy się, że opisanie ich za pomocą innych miar (np. mediany zamiast średniej) nie przynosi innych niż forsowane przez nas rezultatów (przykładem tego rodzaju konfuzji mogą być spory dotyczące użyteczności i rzetelności publikowanych przez GUS danych o przeciętnym wynagrodzeniu w gospodarce narodowej).

---

41 Program Kontroli Jakości Pracy Ankierów.

Stawiamy pytania dotyczące wiarygodności naszych obliczeń – możemy być pewni, że jeśli my tego nie zrobimy, prędzej czy później zrobią to nasi oponenci.

### **7.2.6. Grzech szósty: promowanie „dziwacznych” wyników**

Nie wszystkie mankamenty danych są widoczne na pierwszy rzut oka. Nie wszystkie wynikają też z ewidentnych i łatwych do zidentyfikowania błędów. Musimy jednak pamiętać, że korzystanie z danych, których nie jesteśmy w stanie uwiarygodnić, naraża naszą kampanię i naszą organizację na duże ryzyko (by przypomnieć wskazywany już wyżej przykład kampanii Fundacji „Maciuś”). Dlatego warto zwrócić uwagę, czy przeprowadzone przez nas badania lub analizy są zgodne z innymi dostępnymi oszacowaniami dotyczącymi podobnej problematyki, a także – jeśli to możliwe – czy dają wyniki spójne w czasie.

Jeśli tak nie jest, a mimo to mamy przekonanie o rzetelności naszych wyliczeń, musimy przygotować przekonującą argumentację, która pozwoli obronić je w debacie, jaką zamierzamy wywołać. Możemy być pewni, że prędzej czy później inne strony dyskusji zwrócą uwagę na niekonsekwencje w naszych danych lub odniosą się do innych, korzystniejszych z ich punktu widzenia wyników, dążąc do zdyskredytowania naszego stanowiska. Jeśli nie jesteśmy w stanie wypracować takiej argumentacji lub nie wiemy dobrze, skąd wzięły się rozbieżności w naszych wynikach, najlepiej z nich nie korzystać.

### **7.2.7. Grzech siódmy: nadinterpretacja**

Nawet najlepsze i najbardziej precyzyjne dane nic nie znaczą bez interpretacji. Umiejętność formułowania właściwych wniosków i zachowania przy tym odpowiedniej równowagi pomiędzy tym, co chcemy przekazać, a tym, co wynika bezpośrednio z danych jest kluczem do udanej kampanii. Możliwe (i najczęstsze) są tu zarówno błędy polegające na nadinterpretacji zgromadzonego materiału i podporządkowywaniu go pewnej z góry założonej „tezie”, jak i sytuacja przeciwna – kiedy nasze wnioski są zbyt zachowawcze lub wręcz sprowadzają się do czysto opisowej prezentacji danych. W tym drugim przypadku ryzykujemy jedynie tym, iż odbiorcy naszego komunikatu nie odczytają przesłania kryjącego się w naszych danych. W przypadku pierwszym ryzykujemy utratę wiarygodności.

Do szczególnie częstych błędów tego rodzaju należą te związane z projektowaniem związków przyczynowo-skutkowych w odniesieniu do zależności, które nie muszą wcale mieć takiego charakteru, lub wręcz mogą mieć charakter pozorny.

Przykładem takiej sytuacji może być kampania zmierzająca do oswojenia rodziców z ideą posyłania ich pociech do szkół w wieku 6 lat. Chcąc przekonać do tego pomysłu, Minister Edukacji Narodowej wystąpiła z przekazem, którego główna teza brzmiała: „90% rodziców, którzy zdecydowali się na ten krok, uważa to za dobrą decyzję”. Argumentacja ta wywołała gwałtowną reakcję oponentów reformy. Wskazywali oni m.in. na to, iż deklaracje zadowolenia z wcześniejszego posłania dzieci do szkół siłą rzeczy dotyczyły tylko tych rodziców, którzy taką właśnie decyzję podjęli (a więc uznali, że ich dzieci sobie w szkole poradzą). Innymi słowy, zakwestionowany został sugerowany w kampanii przyczynowo-skutkowy związek między zgodą na wcześniejszą edukację dzieci a satysfakcją z wcześniejszego ich nauczania.



## 8. Wizualizacja danych

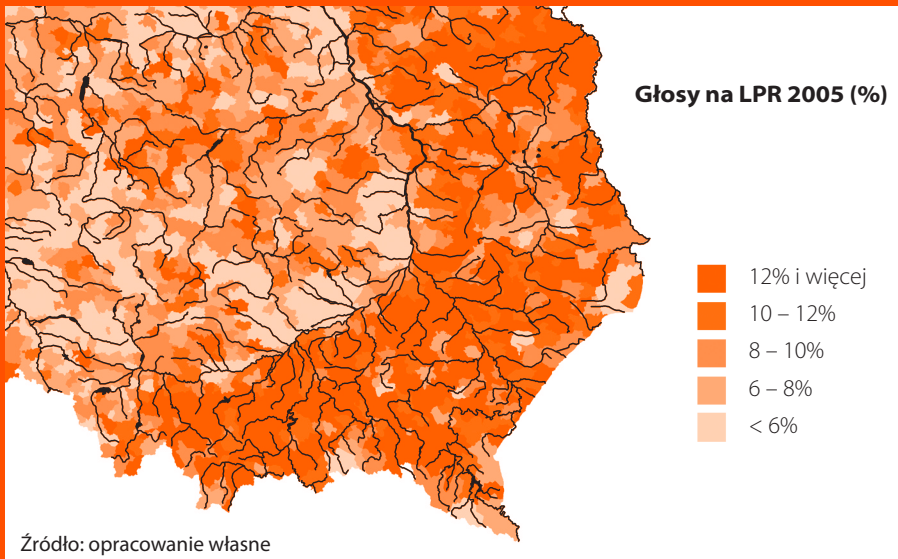
Wizualizacja danych jest często ważnym elementem ich analizy. Niektóre cechy pozostają dla nas nieuchwytnie lub trudne do interpretacji, jeśli nie zobaczymy ich w postaci odpowiedniego rodzaju wykresu. Same analizy tabelaryczne nie pozwolą nam na przykład dostrzec, że podział głosów oddanych na Ligę Polskich Rodzin w wyborach parlamentarnych w 2005 roku przebiegał... wzdłuż linii Wisły! Bez zrobienia wykresu rozkładu wyników matury w roku 2010 nie zauważymy, że część uczniów o wynikach bliskich zaliczającego minimum była ewidentnie podciągana przez egzaminatorów! **[obrazy 8.1. i 8.2. str. 98]**

Jednak w działaniach rzeczniczych graficzna prezentacja danych jest znacznie ważniejsza niż w typowych zastosowaniach analitycznych. W dzisiejszych czasach obrazy to podstawowa forma komunikacji z odbiorcami naszych kampanii, tym bardziej, że często kierujemy je do ludzi, którzy nie są przygotowani do analizy rzędów cyfr czy wyliczeń lub nie mają na to czasu. Trafiony, pobudzający do myślenia obraz jest dziś wart więcej niż tysiąc słów. Zwykle wykresy przestają wystarczać – chcemy, żeby były to małe dzieła sztuki, nie tylko treściwe, ale pobudzające do myślenia i... po prostu ładne.

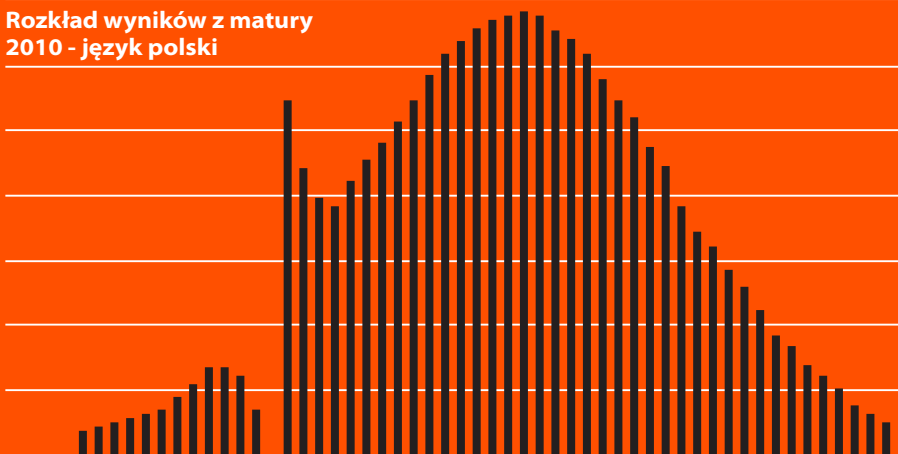
Świadomość tego faktu przekłada się dziś na rozkwit całego „przemysłu” komunikacji wizualnej. Niekiedy wydaje się wręcz, że rozwój ten osiąga powoli punkt, poza którym zaczyna się jego powolna degeneracja. Coraz więcej jest przykładów infografik i wizualizacji, które przestają cokolwiek wyjaśniać, a stają się sztuką dla sztuki albo świadomą manipulacją. Dążenie do uproszczenia przekazu, stanowiące punkt wyjścia całej idei, ustępuje wyścigowi na interaktywność, animację i trójwymiarowość, a forma przerasta treść.

Komunikowanie danych poprzez obrazy to kluczowa broń w działalności rzeczniczej. Warto jednak pamiętać, że może być to broń obosieczna. To, co atrakcyjne w wizualizacji, nie zawsze musi być tożsame z tym, co dobre, a nadmiernie wymyślne, przerysowane, przeładowane efektami czy po prostu błędne grafy mogą przykryć zawarty w nich przekaz lub go zniekształcić. Dlatego zanim przystąpimy do opracowywania czy zamawiania infografik, poznajmy podstawowe zasady wizualizacji stosowane na gruncie analizy danych. Ich świadomość uczyni nasz przekaz bardziej profesjonalnym i pozwoli uniknąć opisanych wyżej pułapek.

Ostatecznie znaczenie takich zasad może być różne w zależności od tego, kto jest odbiorcą danego komunikatu. Poniżej przedstawiamy tylko te, które wydają nam się mieć względnie



### Rozkład wyników z matury 2010 - język polski



Źródło: CKE, Osiągnięcia maturzystów w 2010 r. ([http://www.cke.edu.pl/images/stories/001\\_Matura/WYNIKI/raport\\_matura\\_2010.pdf](http://www.cke.edu.pl/images/stories/001_Matura/WYNIKI/raport_matura_2010.pdf))

**Obrazy 8.1. i 8.2. Wizualizacja jako element analizy danych**

uniwersalny charakter. Traktujemy je jako swojego rodzaju abecadło wizualizacji: to reguły, o których warto pamiętać, kiedy zaczynamy „przemawiać obrazami”.

Istnieje sporo gotowych i szeroko dostępnych zestawień zasad wizualizacji danych. W Internecie stosunkowo łatwo można znaleźć blogi i serwisy poświęcone tej tematyce (większość z nich jest w języku angielskim – zobacz np. FlowingData (<http://flowingdata.com>), JunkCharts (<http://junkcharts.typepad.com>), Chandoo (<http://chandoo.org>). Strony te oferują nie tylko wskazówki dotyczące złotych standardów wizualizacji, ale także przykłady typowych złych praktyk w tej dziedzinie (niekiedy bardzo zabawne), a także odniesienia do pomocnych, często darmowych narzędzi umożliwiających stworzenie własnych wykresów lub grafik (niektóre z nich omawiamy pod koniec rozdziału). Nie dysponując miejscem, by opisać je wszystkie, zwrócimy niżej uwagę na te normy, które są powszechnie podzielane i wydają nam się kluczowe. Każdą z nich opiszemy, przedstawiając powody, dla których dana zasada jest uznawana za ważną, oraz przykłady błędów wynikających z jej nieprzestrzegania.

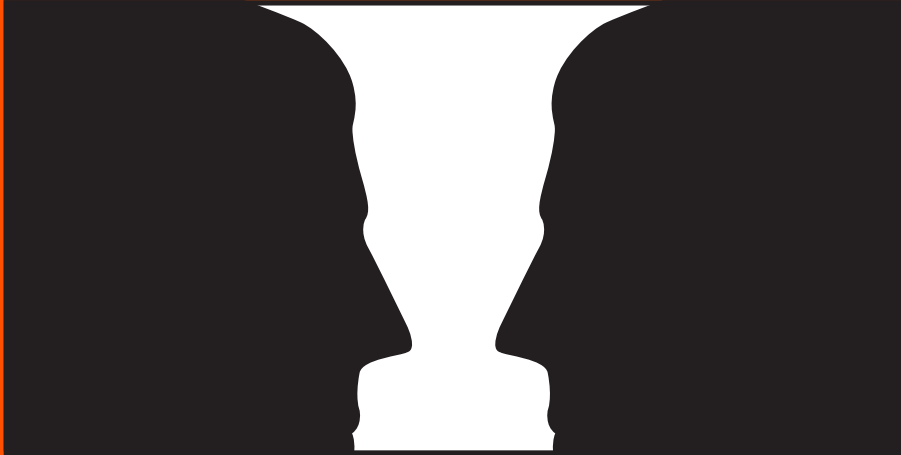
## 8.1. Punkt wyjścia: jak widzimy?

Zanim zaczniemy mówić o konkretnych zasadach, uświadommy sobie, iż wzrok jest jednym z najbardziej „upośledzonych” zmysłów człowieka. Badania neurologiczne i optyczne już pod koniec XIX wieku wskazywały, że z czysto optycznego punktu widzenia oko ludzkie jest bardzo niedoskonałym źródłem informacji. Korzystamy z niego, dokonując wielu założeń uzupełniających niekompletne i niedoskonałe dane, które uzyskujemy za jego pośrednictwem. Krótko mówiąc, nie tyle „widzimy”, co „postrzegamy” rzeczywistość, korzystając z tego, czego się nauczyliśmy jako niemowlęta (warto przypomnieć, że zarówno na początku, jak i pod koniec życia, często tracimy zdolność widzenia nie tylko ze względu na problemy z oczami, ale ze względu na procesy starzenia zachodzące w naszym mózgu). Interpretując obrazy rejestrowane przez siatkówkę, zakładamy m.in., że światło pada z góry, obiekty nie są zwykle widziane od dołu, twarze rozpoznajemy w pozycji pionowej, najpierw koncentrujemy wzrok na obiektach, które są bliżej itp.

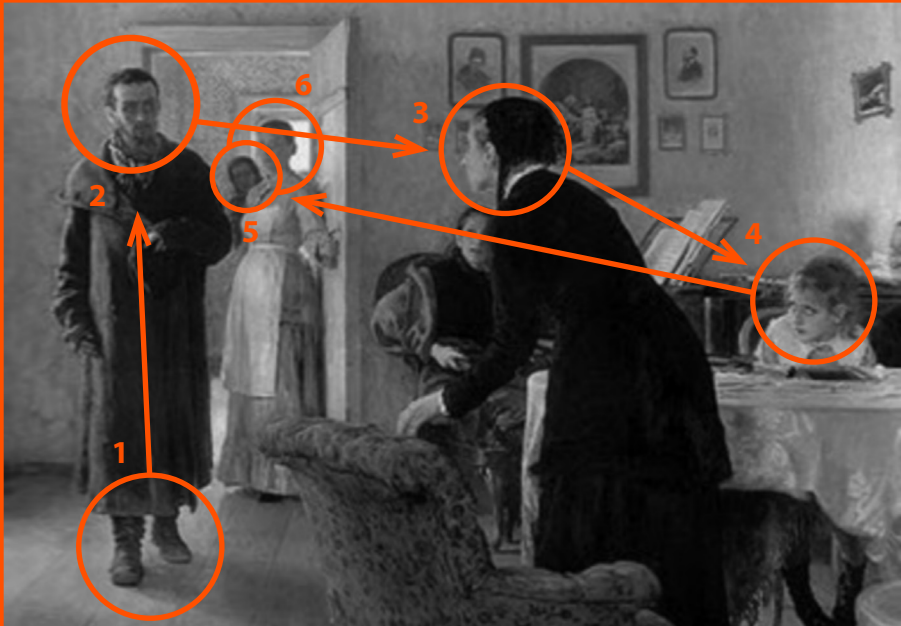
Na te nawykowe odruchy nakładają się cechy wynikające z fizycznych ograniczeń ludzkiego zmysłu wzroku, a także atawizmy. Do tych ostatnich należy na przykład większa wrażliwość wzroku na ruch niż na fakturę, zwiększona wrażliwość na ruch na granicy pola widzenia, ogniskowanie wzroku na elementach o czerwonej barwie, koncentrowanie wzroku na twarzach i wyszukiwanie ich w polu widzenia, organizowanie obrazu w znane nam bryły i krawędzie

itp. **[patrz obrazy 8.3 i 8.4 str. 101]**. Do ograniczeń fizycznych zaliczyć można iluzje związane z nieprawidłowym odbiorem przez wzrok wzorów o różnej jasności, kolorze, kontraście i perspektywie. Receptory nadmiernie pobudzone przez dany obraz mogą na przykład wykazywać aktywność nawet wtedy, kiedy obraz zniknie (tak widzimy tzw. „obrazy widmowe” – **[patrz obraz 8.5 str. 102]**). Często padamy też ofiarą tzw. „hamowania obocznego”, czyli procesu, w którym receptor pobudzony przez wysokie natężenie światła albo inny bodziec zakłóca działanie innych pobliskich receptorów, wskutek czego mamy wrażenie, że pewne elementy obrazu (krawędzie, przecięcia) mają inny kolor lub kształt niż w rzeczywistości **[patrz obraz 8.6 str. 102]**. Nasz wzrok jest także słabo przystosowany do odczytywania perspektywy i analizy obrazów trójwymiarowych **[patrz obrazy 8.7 i 8.8 str. 103]**. Mamy tendencję do interpretowania kolorów poprzez stosunek do tła, oceniamy nachylenie i kąty poprzez odniesienie do kontekstu itp. **[patrz obrazy 8.9, 8.10, 8.11 i 8.12 str. 104 i 105]**. Nasz mózg dąży także do łączenia ich w znane sobie i sugerowane przez kontekst całości **[patrz obraz 8.13 str. 106]**. Może to prowadzić m.in. do sytuacji, kiedy określone obiekty dwuwymiarowe postrzegamy jako bryły (obiekty trójwymiarowe), które w rzeczywistości nie mogą istnieć (struktury takie nazywane są „obiettami niemożliwymi” – **[patrz obraz 8.14 str. 106]**).

Niedoskonałości ludzkiego wzroku jest znacznie więcej – by wspomnieć choćby o problemach z właściwym odwzorowaniem czy rozróżnianiem kolorów, oceną odległości lub rozmiarów elementów itp. Niezależnie od nich, o naszej zdolności do właściwej interpretacji obrazu decydują oczywiście kwestie niezwiązane z samym postrzeganiem – kompetencje, inteligencja, spostrzegawczość, przeświadczenia. Wszystkie te czynniki znajdują odzwierciedlenie w sposobie, w jaki odbieramy graficzne reprezentacje danych. Można z nich świadomie skorzystać, kiedy chcemy przyciągnąć uwagę obserwatora (lub go oszukać), ale także można się na nich zawieść, jeśli usiłujemy uraczyć go zbyt dużą ilością niepotrzebnych fajerwerków. Pamiętajmy więc o tym, że obraz, który tworzymy, ma być formą jednoznacznego, wymownego komunikatu dla odbiorcy – historią w pigułce. Nie może być ani zbyt przeładowany informacją, ani pozbawiony sensu. Przede wszystkim jednak, powinien być tak skonstruowany, aby nie wprowadzać odbiorcy w błąd. Chcąc nad tym zapanować, warto przestrzegać kilku złotych reguł wizualizacji.



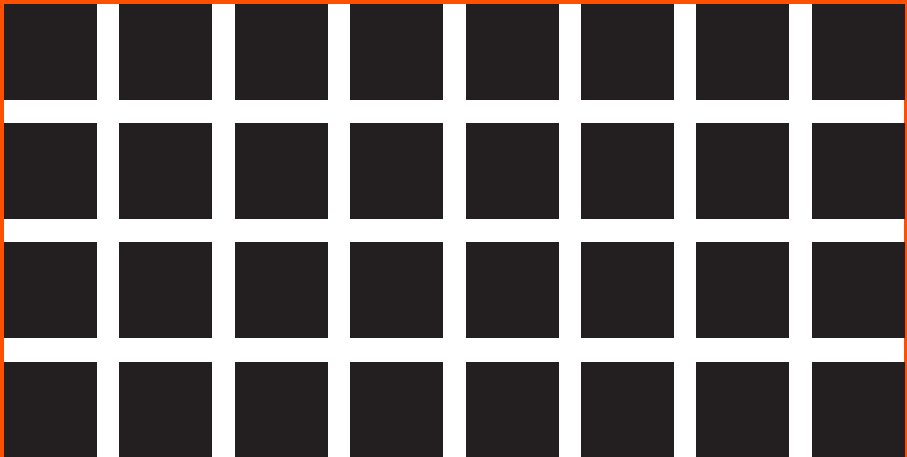
Obraz 8.3. Twarze czy waza?



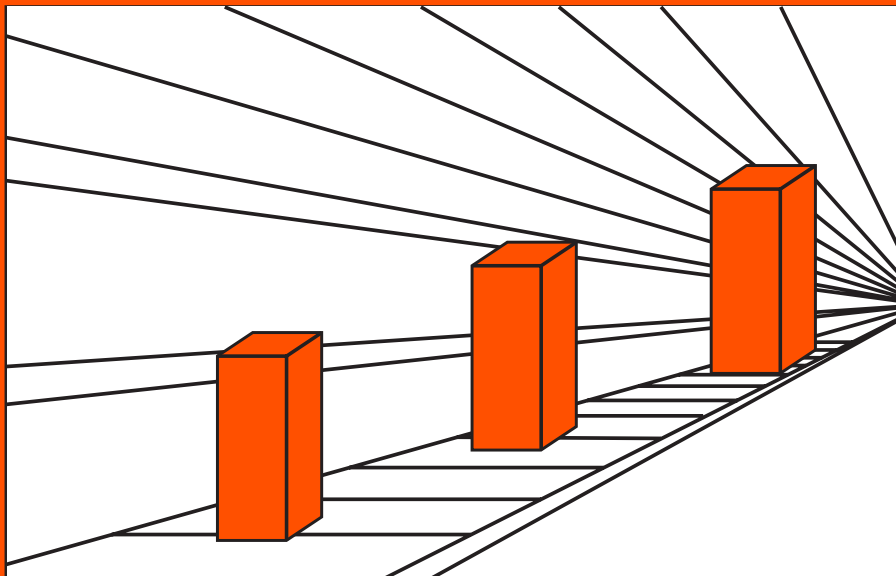
Obraz 8.4. Studium ruchów oka A. Yarbusa (1967).

# 30 sekund

Obraz 8.5. Obrazy widmowe: patrz na obraz przez 30 sekund, potem spójrz na białe tło.



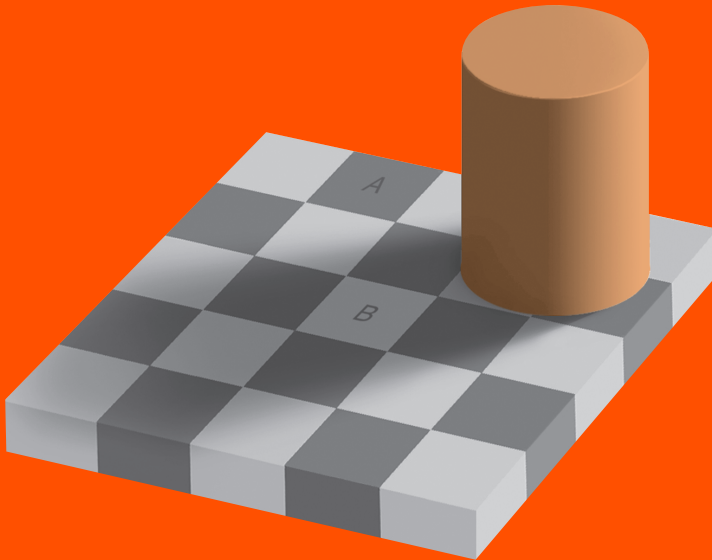
Obraz 8.6. Skutki „rywalizacji” między receptorami w naszym oku – na przecięciach linii widzimy ciemne punkty.



Obraz 8.7. Perspektywa – trudno nam precyzyjnie ocenić odległość słupków i ich wysokość.



Obraz 8.8. Tzw. Złudzenie Ponza. Wbrew pozorom, obie żółte kreski są tej samej długości.

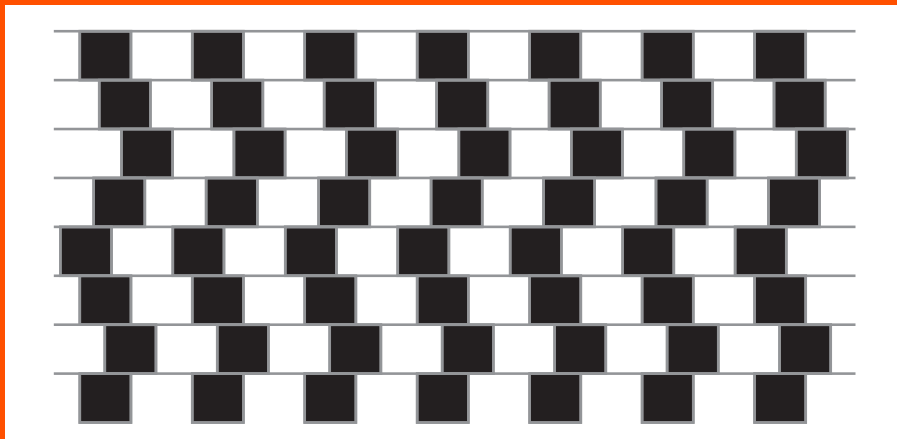


Obraz 8.9. Ocena kolorów poprzez odniesienie do pobliskiego tła – elementy A i B mają w rzeczywistości ten sam kolor.

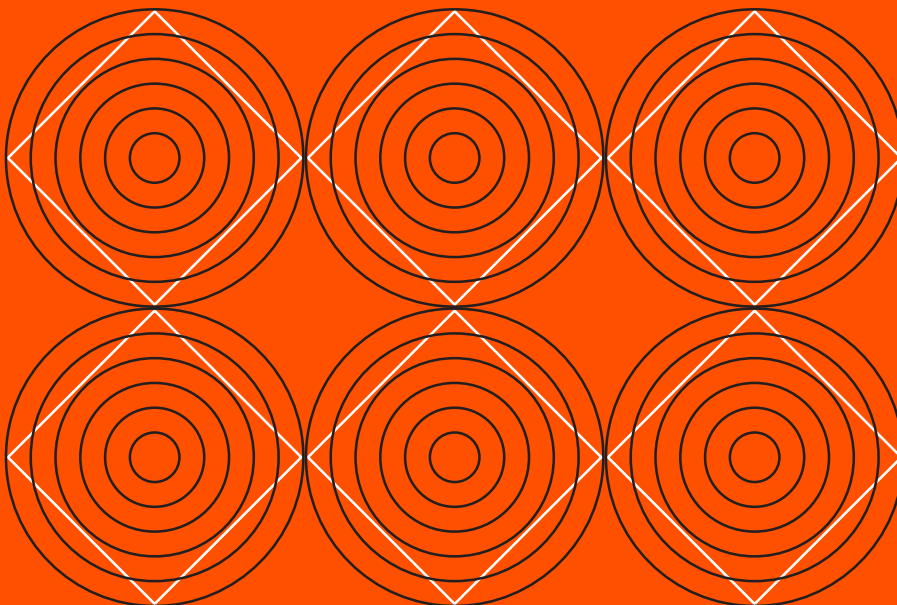


Obraz 8.10. Iluzja kontrastu – tło jest gradientowe. Jego kolor wywołuje iluzję gradientowego wypełnienia paska pośrodku – w rzeczywistości pasek jest w jednym kolorze.





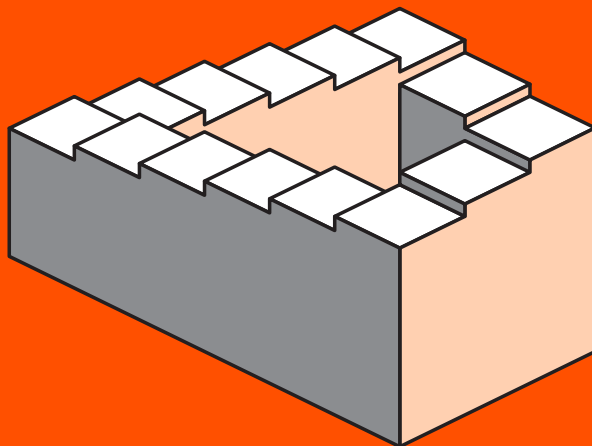
Obraz 8.11. Poprzesuwane białe i czarne kwadraty wywołują złudzenie nachylenia linii. W rzeczywistości są one równoległe.



Obraz 8.12. Promieniste tło „zakrzywia” linie biegnące po przekątnej. W rzeczywistości są to linie proste.



Obraz 8.13 Trójkąt Kanizsa: nasz mózg „wytwarza” figurę na podstawie kilku elementów.



Obraz 8.14. Schody Penrose'a – zawsze wyżej.

Źródło: Wikipedia

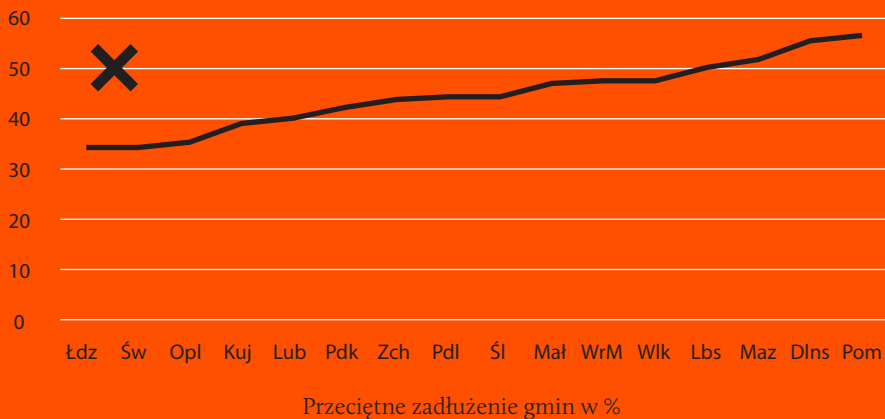
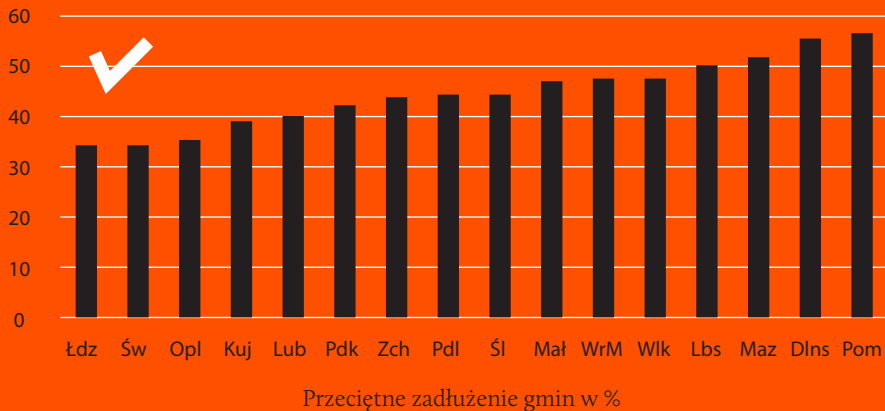
## 8.2. Zasady wizualizacji

### 8.2.1. Zasada 1. Nie pokazuj faktów, których nie ma

Bardzo często natrafiamy na wizualizacje sugerujące istnienie faktów, które w istocie nie miały miejsca. Nie zawsze musi to być świadoma manipulacja (do takich sytuacji jeszcze niżej wrócimy) – najczęściej przyczyną jest nieuwaga, niekompetencja czy brak przezorności autorów. Tego rodzaju błędem jest np. niewłaściwy dobór typu wykresu ze względu na charakter prezentowanych danych. Jeśli prezentujemy dane dotyczące natężenia pewnej cechy w grupach, które nie tworzą żadnego naturalnego porządku (nie da się tych grup uporządkować od najmniejszej do największej), nie stosujemy wykresów liniowych. Linia sugeruje bowiem występowanie przedstawianego na wykresie zjawiska także pomiędzy kategoriami na osi, a co za tym idzie – pewną ciągłość czy kierunkowość zmian w jego natężeniu. Jeśli np. opisujemy przeciętny stan zadłużenia gmin w poszczególnych regionach Polski, przebieg tej linii pomiędzy punktami oznaczającymi wyniki regionów nic nie oznacza. Wykresy liniowe wykorzystujemy tylko wtedy, kiedy cecha stanowiąca kategorie osi poziomej ma charakter ilościowy i ciągły [[patrz podrozdział 7.1.14. „Zmienne ilościowe”, str. 87](#)] lub gdy mówi o czasie. [[patrz obraz 8.15. str. 108](#)]

Jeśli prezentujemy nasze dane na wykresie liniowym, pamiętajmy, żeby uwzględnić w wizualizacji wszelkie nieciągłości danych. W tego typu wykresie jednostki odkładane na osi poziomej powinny oddzielać równe interwały (np. każda kolejna jednostka to kolejny rok) [[patrz obraz 8.16.A str.109](#)]. Nierzadko jednak zdarza się, że nie dysponujemy danymi dla pełnego zakresu tych jednostek (na przykład mamy dane roczne dla okresu 1995-2012, ale z wyłączeniem lat 1998-1999). W takich sytuacjach nie usuwajmy brakującego roku z osi poziomej (tworzymy wówczas iluzję ciągłości danych i wpływamy na przebieg krzywej na wykresie [[patrz obraz 8.16.D str.109](#)]). Zaznaczmy w jakiś sposób, że część przebiegu linii na wykresie jest tylko naszym oszacowaniem wartości prezentowanej zmiennej [[patrz obraz 8.16.B str.109](#)], albo w ogóle nie rysujemy jej dla tych kategorii osi poziomej, dla których danych nie ma [[patrz obraz 8.16.C str.109](#)].

Sytuacji, kiedy komunikujemy odbiorcy „nieistniejące fakty”, jest więcej. Nie stosujemy wykresów kołowych tam, gdzie dane, które prezentujemy, nie sumują się do 100% lub też po zsumowaniu przekraczają wartość 100%. Tak naprawdę najlepiej będzie, jeśli w ogóle nie będziemy stosować wykresów kołowych [[patrz podrozdział 8.2.6. „Zasada 6”, str. 120](#)]. Bądźmy też czujni, upiększając nasz wykres – robiąc to, możemy nieświadomie zmienić jego znaczenie. Na przykład, tworząc wykres kołowy o niestandardowym, nieforemnym kształcie, uniemożliwiamy

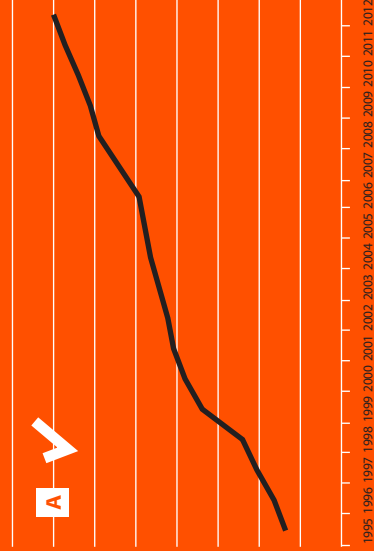


**Obraz 8.15. Niewłaściwe użycie wykresu liniowego<sup>1</sup>**

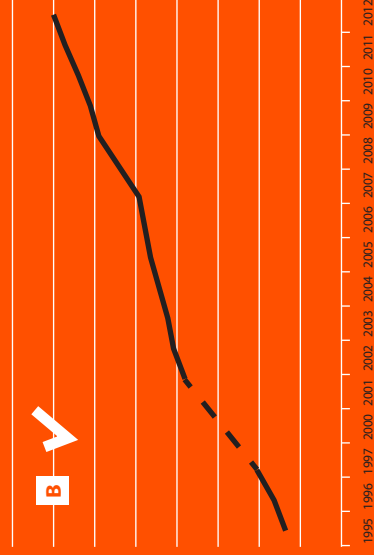
<sup>1</sup> Wykresy nie przedstawiają rzeczywistego poziomu zadłużenia – zostały wygenerowane losowo, wyłącznie dla celów prezentacji.

4 000 zł  
3 500 zł  
3 000 zł  
2 500 zł  
2 000 zł  
1 500 zł  
1 000 zł  
500 zł  
zł

**A** ✓

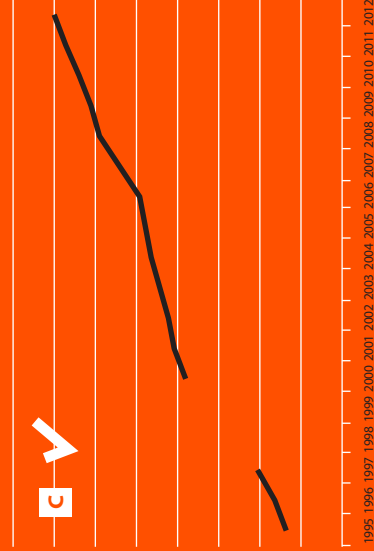


**B** ✓

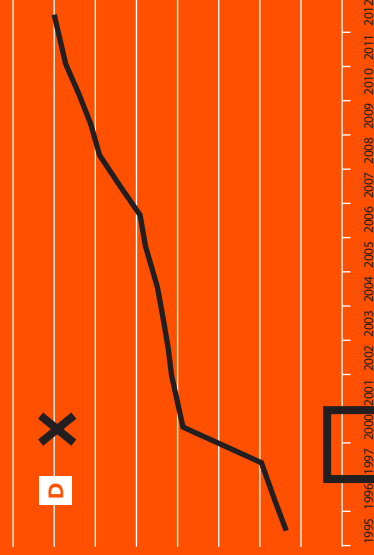


4 000 zł  
3 500 zł  
3 000 zł  
2 500 zł  
2 000 zł  
1 500 zł  
1 000 zł  
500 zł  
zł

**C** ✓

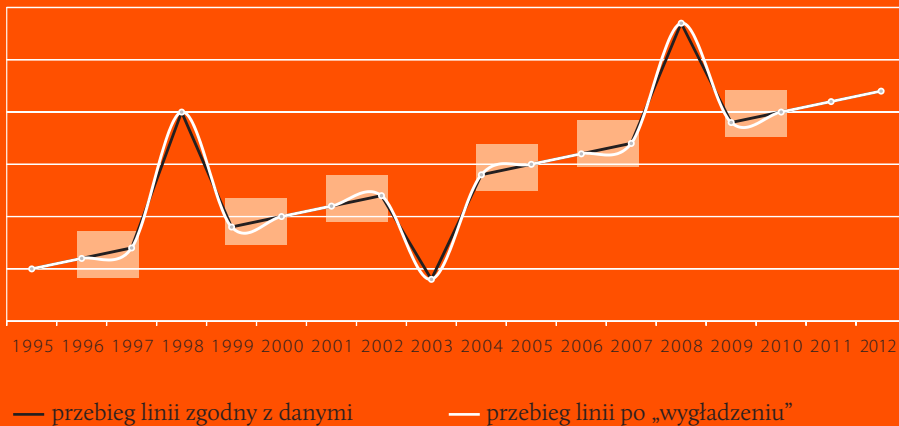


**D** ✗

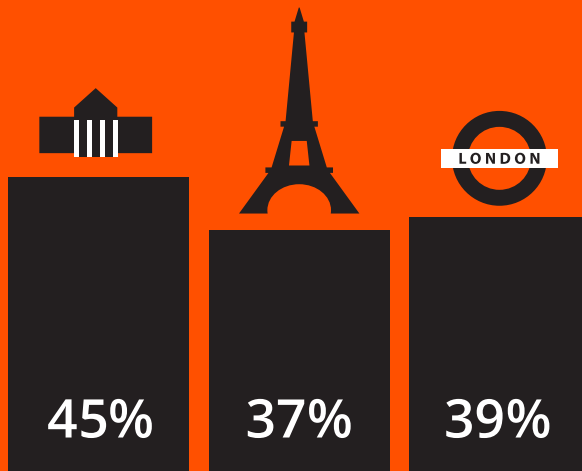


Przeciętnewynagrodzenie w gospodarce narodowej wg GUS

Obraz 8.16. Poprawne i niepoprawne uzupełnianie braków danych na wykresie liniowym



Obraz 8.17. Fałszywe (niewystępujące w danych) zakrzywienia linii wynikające z „wygładzania” jej przebiegu w Excelu (w jaśniejszych polach)



Obraz 8.18. O jeden „gadżet” za daleko: ikony nad słupkami utrudniają interpretację wyników

odbiorcy precyzyjne porównanie prezentowanych na nim kategorii (por. przykład przy zasadzie 6). Wygładzanie linii w Excelu w przypadku, kiedy między dwoma sąsiadującymi kategoriami na osi poziomej mamy do czynienia z gwałtowną zmianą wartości prezentowanej zmiennej, powoduje wyraźne zakrzywienie linii oznaczającej tę zmienną w przeciwną stronę – nieświadomy odbiorca odczyta to jako zmianę wartości zmiennej, choć w istocie za przebieg krzywej odpowiada w tym przypadku wyłącznie błędne działanie oprogramowania [patrz obraz 8.17 str. 110]. Powstrzymajmy się również przed stosowaniem wymyślnych ikon lub logotypów na oznaczenie słupków na naszym wykresie (dotyczy to zarówno umieszczania ich nad słupkami, jak i zastępowania ich jakimikolwiek ikonkami). Mają one różną objętość i kształt, co może prowadzić do błędnej lub co najmniej utrudnionej interpretacji wizualizacji [patrz obraz 8.18 str. 110].

### 8.2.2. Zasada 2. Nie manipuluj

Aby oszukać czytelnika, nie trzeba uciekać się do podawania fałszywych danych czy manipulowania ich dobozem. Niekiedy wystarczy umiejętnie podrasowanie wykresu, na którym przedstawiamy wyniki naszych analiz. Granica pomiędzy wyrazistym przekazem a oszustwem jest w świecie wizualizacji bardzo rozmyta. Jeśli ją przekroczyliśmy (świadomie czy nieświadomie), nie tylko stracimy wiarygodność, ale możemy też narobić sobie wrogów.

#### Które „chwyt” nie są dozwolone?

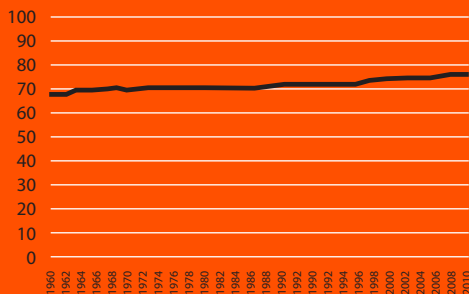
##### Manipulowanie skalą

To jeden z najczęściej wykorzystywanych sposobów na zwodzenie odbiorców analiz. Najbardziej oczywistą formą manipulacji jest w tym wypadku wydłużanie osi Y, a więc ustawianie zbyt dużego w stosunku do prezentowanych danych zakresu wartości przedstawianych na osi. W efekcie uzyskujemy wykres, który jest w znacznym stopniu pusty, a dane, które na nim prezentujemy, wydają się niespecjalnie różnić od siebie [patrz obraz 8.19A str. 112]. Często zdarza się jednak również odwrotna sytuacja – kiedy to zakres skali na osi jest ograniczony tylko do zakresu wartości prezentowanej zmiennej (oś nie zaczyna się od 0). Takie postępowanie jest uzasadnione tylko wtedy, jeśli różnice, które przedstawiamy na wykresie, są godne odnotowania z merytorycznego punktu widzenia [patrz obrazy 8.19.B i 8.20 str. 112].

##### Manipulowanie obrazem

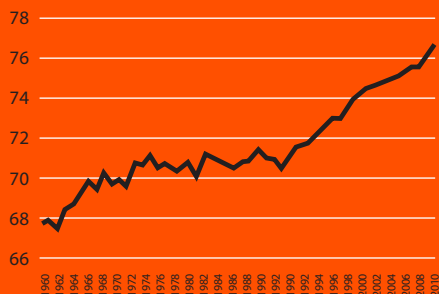
Innym często wykorzystywanym sposobem oddziaływania na odbiorcę jest manipulowanie graficznymi własnościami wizualizacji. Choć w większości przypadków są to wyłącznie zabiegi estetyczne, zdarza się, że służą zniekształceniu obrazu wyłaniającego się z danych. Mogą polegać na dodaniu do wizualizacji tła, które optycznie zmienia (np. poszerza, wydłuża) pewien

A. Zakres skali na osi y zmieniony tak, by ukryć zróżnicowanie

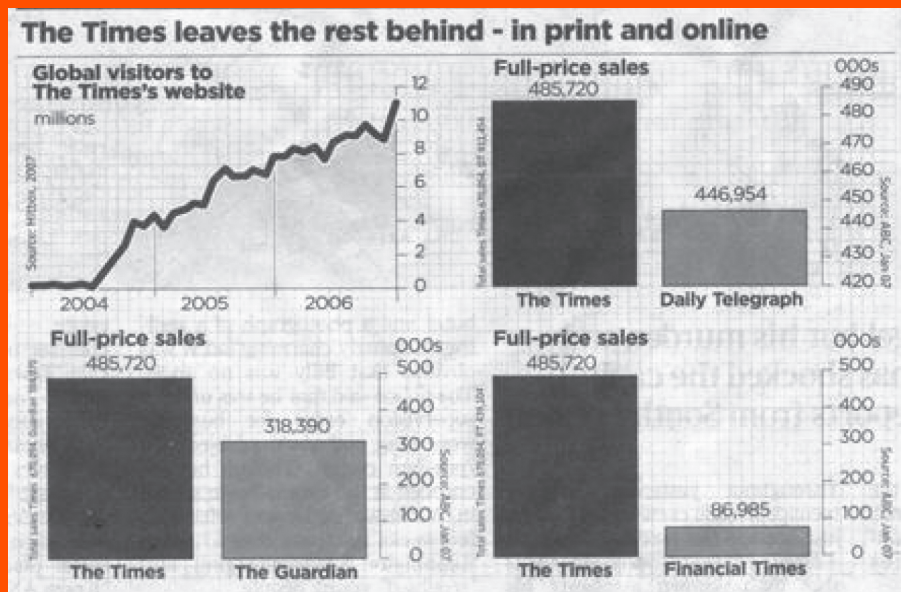


— Przewidywania średnia długość życia przy urodzeniu, Polska

B. Zakres skali na osi y zmieniony tak, by podkreślić zróżnicowanie

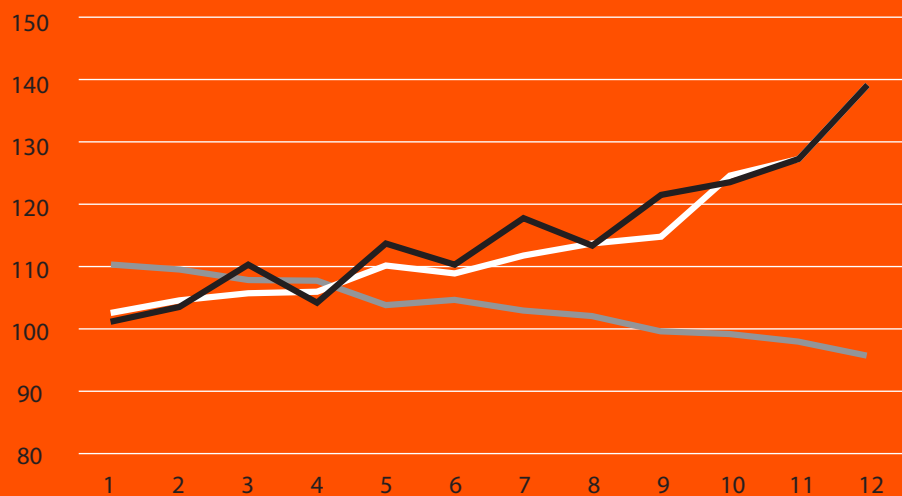
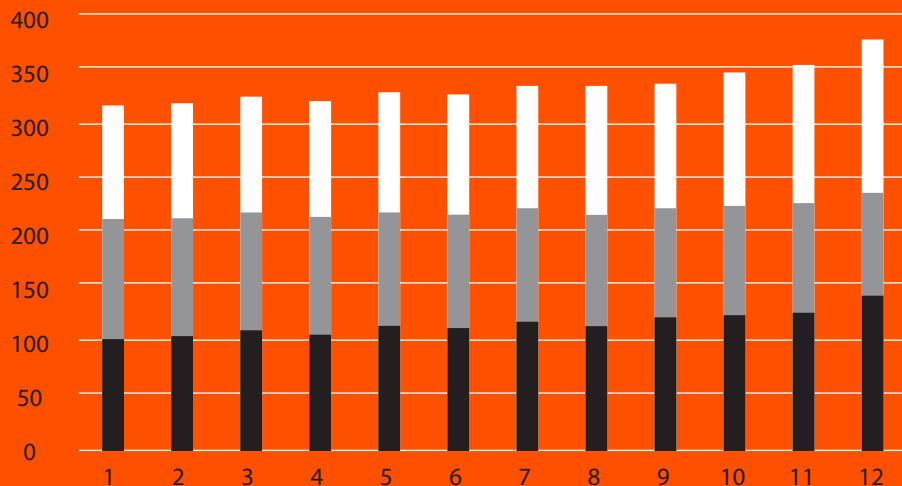


Obraz 8.19 Wykres zmian w przeciętnej długości trwania życia w Polsce od lat 60



Obraz 8.20. Przykład niedopuszczalnej manipulacji skalą: wyniki sprzedaży „The Times” w porównaniu do innych (wybranych) pism. Na wykresie w górnym prawym rogu efekt różnicy między „The Times” i „Daily Telegraph” osiągnięto, ustawiając jako minimum skali wielkość 420 tys. sprzedanych egzemplarzy.





A
  B
  C

**Obraz 8.22. Manipulacja przez dobór rodzaju wykresu – skumulowany wykres słupkowy a wykres liniowy**

fragment wykresu [patrz obraz 8.21.A str. 115], na zniekształceniu proporcji obrazu za pomocą efektów 3D lub perspektywy [patrz obraz 8.21.B str. 115] albo na dodaniu do wykresu rozmaitych artystycznych efektów (animacji, grafik) wpływających na naszą percepcję [patrz obraz 8.21.C str. 115]. Niezależnie od tego, jak ładne są wizualizacje (a najczęściej w ogóle nie są), nie spełniają one podstawowych kryteriów profesjonalnej prezentacji danych. Opublikujmy je, a możemy z dużym prawdopodobieństwem liczyć na miejsce w którymś z coraz liczniejszych blogów piętnujących „złe wykresy”. Oczywiście nie oznacza to, że należy wystrzegać się wszelkich prób uatrakcyjnienia wykresów – w końcu cała sztuka wizualizacji służy temu, aby za ich pomocą oddziaływać na odbiorcę. Proponujemy jednak regułę: nie korzystajmy z opcji graficznych, które w jakikolwiek sposób utrudniają precyzyjną interpretację prezentowanego materiału.

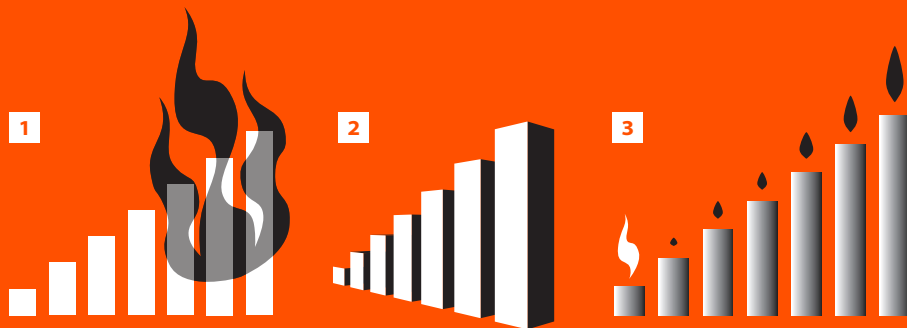
### **Manipulowanie poprzez dobór rodzaju wykresu**

Nie musimy dopuszczać się jawnej manipulacji obrazem, aby wpłynąć na jego odbiór – czasami wystarczy umiejętny dobór rodzaju wykresu. Jeśli na przykład analityk nie chce, aby czytelnik szczegółowo porównywał różnice między kategoriami na schemacie, prezentuje dane na dwóch wykresach kołowych, do tego o różnej wielkości [patrz podrozdział 8.2.6 „Zasada 6”, str. 120]. Jeśli chce pokazać wzrost wartości jakiejś cechy, nawet jeśli wzrost ten wcale nie jest permanentny, albo wręcz ukryć spadek wartości tej cechy, może przedstawić ją na wykresie słupkowym skumulowanym. Podstawa poszczególnych słupków dla tej cechy (poza słupkiem na samym dole) jest zmienna, co skutecznie utrudnia precyzyjną interpretację wykresu.

Formą manipulacji jest też **nierównomierna agregacja danych w kategorii**, dla których prezentowane są dane. Dotyczy to sytuacji, kiedy prezentujemy dane o liczebności, udziale lub wartości w grupach wyróżnianych przez przynależność do przedziału jakiejś cechy. Odpowiedni ich dobór może pomóc wyróżnić określoną kategorię spośród innych – na przykład udowodnić, że jest ona wyjątkowo liczna lub istotna w kategoriach wielkości prezentowanej zmiennej.

### **8.2.3. Zasada 3. Proste jest ładne**

To, jak wygląda nasza wizualizacja, zależy oczywiście od tego, do kogo ją kierujemy. Dla celów kampanii społecznej będziemy uciekać się do innych rozwiązań niż wtedy, kiedy przygotowujemy naukowy raport. Namawiamy jednak do tego, aby zawsze dążyć do maksymalnego uproszczenia wizualizacji, zgodnie z zasadą: proste jest piękne. W czasach, kiedy niemal każde oprogramowanie służące wizualizacji danych oferuje wiele możliwości skomplikowania spraw, prostota przekazu zaczyna stanowić największą wartość. Zatem o ile nie jest to z jakichś powodów niezbędne, proponujemy:

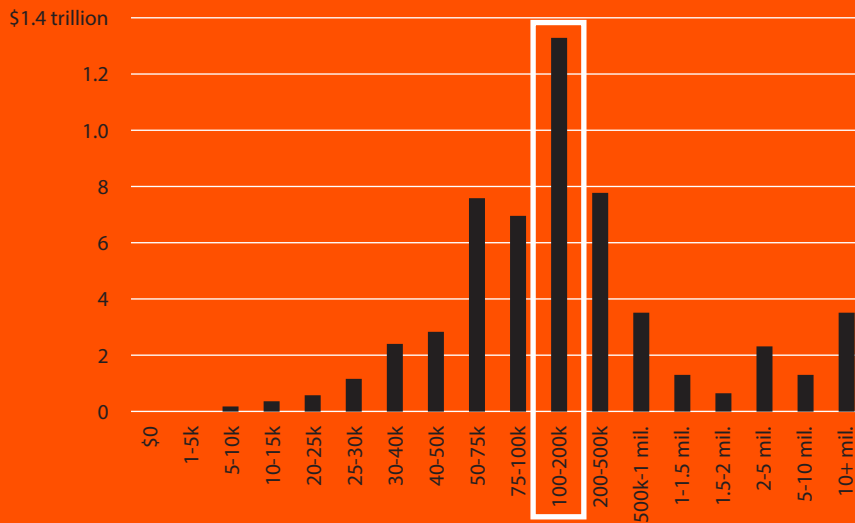


**Obraz 8.21. Graficzne „fajerwerki” utrudniające interpretację wykresu.**

1. Tłó; 2. B. Perspektywa i efekty 3D 3. Grafika.

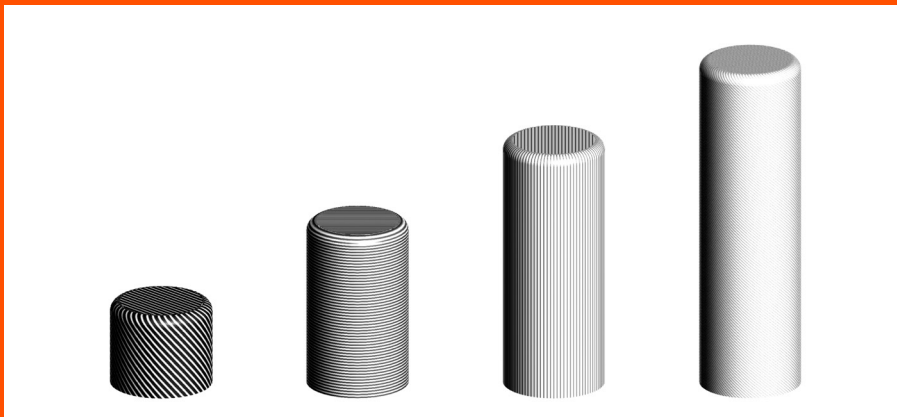
### The Middle Class Tax Target

The amount of total taxable income (left scale) for all filers by adjusted gross income level for 2008



**Obraz 8.23. Manipulowanie grupowaniem danych w kategorii**

Źródło: „The Wall Street Journal”

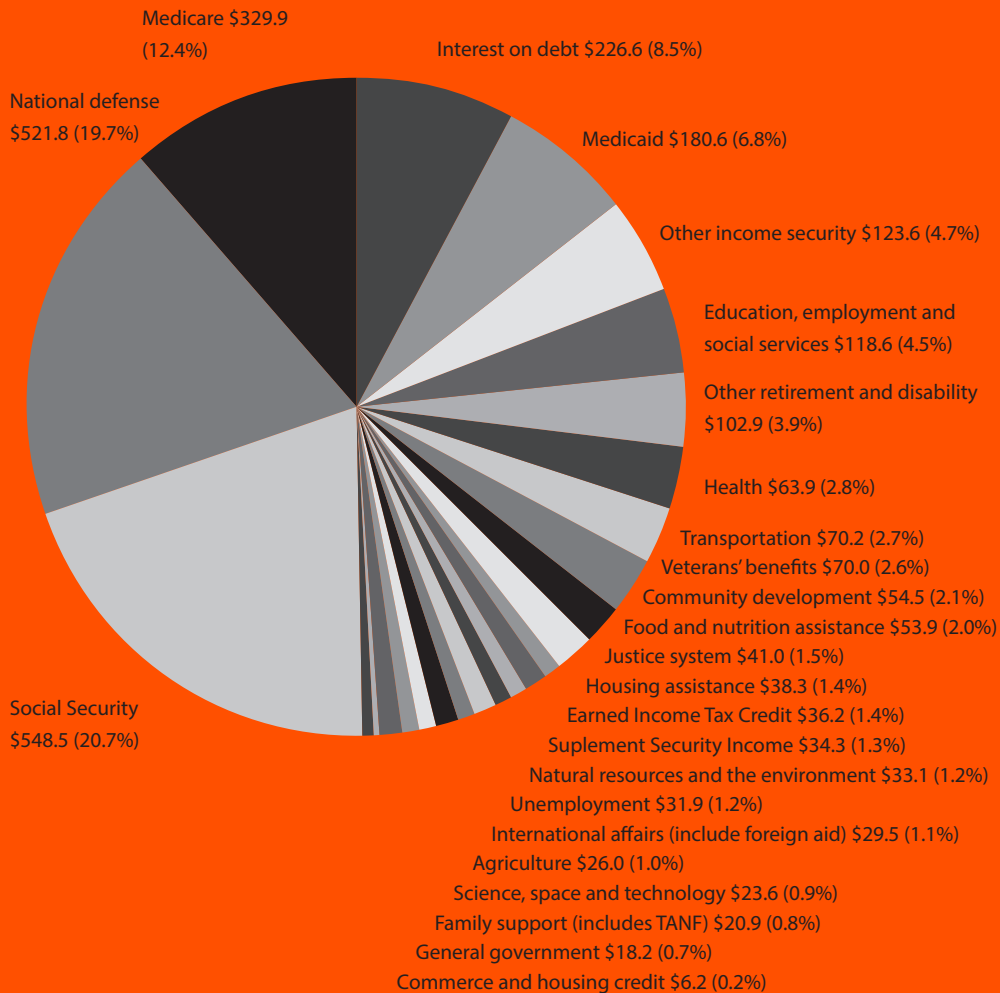


Obraz 8.24. Zróżnicowane tekstury odciągają naszą uwagę do ważnych elementów wykresu



Obraz 8.25. Tło, które skutecznie konkuruje z danymi

Źródło: opracowanie własne



Źródło: Blog amerykańskiej organizacji farmaceutów ASPET

**Obraz 8.26. Za dużo informacji: jedyne wartościowe źródło informacji to etykiety.**

- unikać błyszczących lub niejednorodnych tekstur **[patrz obraz 8.24 str. 116]**;
- dostosować typ i format wykresu do możliwości odbiorców (np. nie prezentować wykresów rozbudowanych sieci czy tzw. *treemap*, jeśli wiemy, że dla czytelnika nie wynika z nich nic więcej niż świadomość, że potrafimy je opracować);
- zostawić w spokoju tło wykresów i opcje związane z graficznym wyróżnianiem poszczególnych kategorii z wyjątkami, o których piszemy w dalszej części **[patrz niżej podrozdział 8.2.5. „Zasada 5”] [patrz obraz 8.25 str. 116]**;
- zrezygnować z efektów 3D **[patrz podrozdział 8.2.2. „Zasada 2” str. 111]**.

#### 8.2.4. Zasada 4. Lepiej mniej, ale lepiej

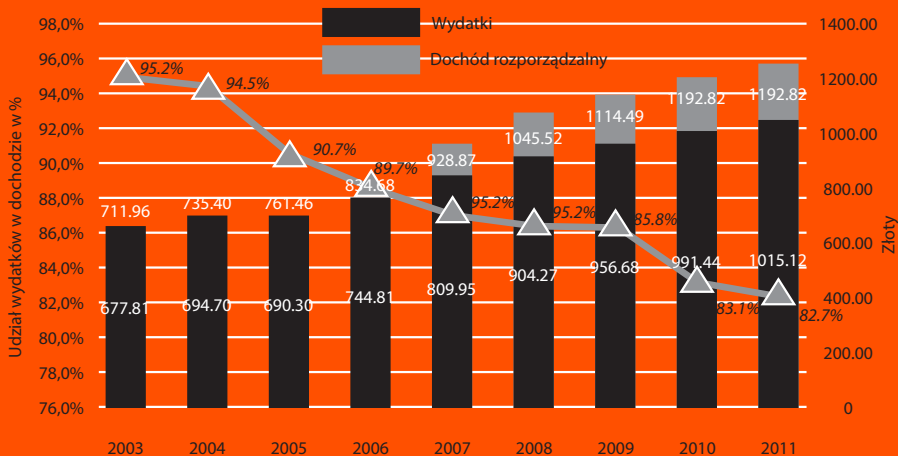
Słynne hasło Lenina w dziedzinie wizualizacji danych sprowadza się do postulatu, aby nie umieszczać na wykresie zbyt dużej ilości (konkurujących ze sobą) informacji. Nie prezentujmy więc na jednym schemacie zbyt pokażnej liczby kategorii lub elementów (por. obraz 8.26). Zamiast tego, wybierzmy kilka takich, na których chcemy skoncentrować komunikat. Powstrzymajmy się też, o ile to możliwe, przed umieszczaniem na nim zbyt dużej liczby zmiennych (zwłaszcza przed stosowaniem tzw. osi pomocniczych) oraz nadmiernie szczegółowych etykiet danych **[patrz obraz 8.27 str. 118]**.

#### 8.2.5. Zasada 5. Używaj kolorów z głowy

Jak już się przekonaliśmy, ludzkie oko nie jest dobrze przystosowane do odbioru niuansów kolorystycznych – to jeden z elementów graficznych, które wyjątkowo często stają się źródłem pomyłek interpretacyjnych. Niektórzy z nas (według niektórych badań nawet do 8% populacji) cierpią na tzw. ślepotę barw, która uniemożliwia lub zaburza rejestrowanie różnic między pewnymi barwami (by wspomnieć tylko najpopularniejsze odmiany: deuteranopia to niezdolność do rozróżniania barwy zielonej, protanopia to problem z odróżnianiem barwy czerwonej, tritanopia – problem z odróżnianiem żółci i niebieskości. Nie wiemy też, jak będą wykorzystywane nasze wizualizacje – czy na przykład ktoś nie postanowi wydrukować ich na czarno-białej drukarce. Jeśli chcemy, aby nasz komunikat był zrozumiały dla wszystkich i w każdych warunkach, warto przestrzegać pewnych zasad.

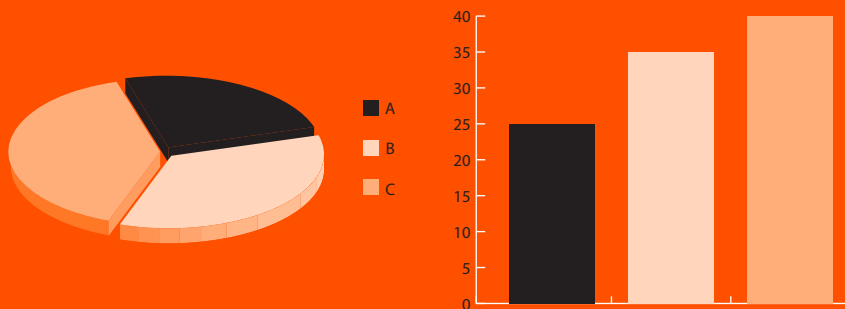
Po pierwsze, proponujemy nie stosować nazbyt pstrokatych kolorów, a zwłaszcza nie używać ich w dużej ilości – niemal na pewno utrudni to czytelnikowi interpretację wykresu. Optymalne (choć nie zawsze możliwe, ze względu na złożoność wizualizacji) rozwiązanie polega na zastosowaniu stonowanych barw dla większości elementów wykresu i jednej wyrazistej (np. czerwonej) dla tych elementów, które chcemy szczególnie wyeksponować. Po drugie, nie korzystajmy

Poziom przeciętny miesięcznych dochodów i wydatków na osobę w gospodarstwie domowym oraz udział wydatków w dochodzie rozporządzalnym 2003-2011



Źródło: Sytuacja gospodarstw domowych w 2011 r. w świetle wyników badania budżetów gospodarstw domowych, GUS 2012.

**Obraz 8.27. Graficzna prezentacja danych w jednym z raportów GUS – za dużo informacji oznacza brak informacji**



**Obraz 8.28. Który z elementów na wykresie przyjmuje największą wartość? Wykres kołowy (do tego 3D) nam tego nie pokaże...**

Źródło: opracowanie własne

z dowolnie wybranej przez siebie palety kolorów (ani z tych oferowanych domyślnie przez najpopularniejsze programy służące do robienia wykresów), ale z takiej, która czytelnie odróżnia kolorowane elementy i do tego sprawdza się w wydrukach czarno-białych. Nie musimy zdawać się przy tym na eksperymenty – wystarczy zastosować jedną z wielu predefiniowanych palet dostępnych w Internecie (por. np. <http://colorbrewer2.org/>). Po trzecie, nie używajmy kolorów gradientowych – zwłaszcza w tle wykresu. I po czwarte (a być może najważniejsze) starajmy się nie różnicować kolorów, jeśli nie jest to potrzebne. Jeśli np. tworzymy sekwencję wykresów, w której wielokrotnie występuje ważna dla nas zmienna – np. wyniki w nauce – starajmy się konsekwentnie oznaczać ją w możliwe ten sam sposób (np. tym samym kolorem).

### 8.2.6. Zasada 6. Stawiaj na precyzję

Wizualizacja to forma wypowiedzi. Tak samo jak nie lubimy okrągłych, niewiele mówiących, pozbawionych treści komunikatów werbalnych, nie polubimy niejasnych, nieprecyzyjnych czy mylących wizualizacji – chyba że skoncentrujemy się wyłącznie na ich stronie estetycznej. Z tego punktu widzenia może zadziwiać popularność różnych form, które nie są dobrze dostosowane do deficytów zmysłu wzroku i łatwo mogą wyprowadzić nas na manowce.

Nasze decyzje dotyczące formy wizualizacji mają duży wpływ na precyzję, z jaką będzie w stanie ocenić je ludzkie oko (czyli nasz odbiorca). Interesujące uszeregowanie możliwości, przed którymi stoimy, przywołali autorzy raportu *Principles of Information Visualization*, Institute for Informatics & Digital Innovation (Kennedy, 2012)<sup>42</sup>. Jak z niego wynika, najlepszym (najbardziej precyzyjnym, pozostawiającym najmniejsze pole wątpliwości) sposobem na prezentację różnic w danych jest odwołanie się do ich pozycji w przestrzeni (szeregowanie, odkładanie jako punktów na osi X, Y). Tam, gdzie to możliwe, dobrą strategią jest też prezentacja za pomocą długości elementów na wykresie (np. za pomocą wykresów słupkowych)<sup>43</sup>. Znacznie trudniejsze do przetworzenia przez człowieka są różnice w kątach, nachyleniu, powierzchni lub objętości porównywanych elementów (co często wykorzystywane jest przeciw

---

42 <http://mkweb.bcgsc.ca/vizbi/2012/principles.pdf>

43 Czasami, jeśli dysproporcje między największymi i najmniejszymi wartościami prezentowanej cechy są zbyt duże, może to być jednak utrudnione. W takich przypadkach warto niekiedy odwołać się do skali logarytmicznej – czyli obliczyć logarytm naturalny lub dziesiętny z wartości poszczególnych punktów serii. Spowoduje to spłaszczenie wykresu w taki sposób, że najniższe wartości będą zdecydowanie bliższe wartościom najwyższym. Jednak może też wprowadzić naszych odbiorców w błąd, zwłaszcza jeśli nie mają odpowiedniego doświadczenia w analizie danych. Kierując się wyłącznie wzrokową oceną różnic między poszczególnymi elementami na wykresie, mogą mylnie odczytać jego przesłanie. Do tego nie będą oni w stanie zinterpretować wartości na zlogarytmizowanej osi.



odbiorcy). Najmniej precyzyjnym narzędziem jest różnicowanie gęstości (czasami stosowane np. w prezentacjach danych na mapach) lub koloru grafiki ze względu na wartość zmiennej.

Co z tego wynika? Przede wszystkim, unikajmy dwóch często spotykanych sposobów prezentacji danych: wykresów kołowych (objętość, kąty) oraz wykresów 3D (nie wspominając już o wykresach kołowych 3D). Te pierwsze mają przynajmniej tę zaletę, że sugerują istnienie pewnej całości (100%), której strukturę pokazuje wykres (co jednak bywa również ich wadą – gdy wykres kołowy wykorzystywany jest do prezentacji danych, które nie tworzą takiej całości – np. pytań z wieloma odpowiedziami, których sumaryczny udział przekracza 100%). Te drugie (wykresy 3D) nie oferują w gruncie rzeczy żadnej wartości dodanej, uruchamiają za to wiele z omówionych w pierwszej części rozdziału zaburzeń percepcji.

Wykresy kołowe nie pomagają w analizie różnic między kategoriami, zwłaszcza jeśli różnice te są stosunkowo niewielkie. Nie sprawdzają się także, jeśli liczba porównywanych kategorii przekracza sześć (również ze względu na potrzebę używania odróżniających się i czytelnych w druku czarno-białym kolorów). Ponadto bardzo łatwo jest wpłynąć na ich interpretację. Ludzkie oko nie ocenia tak samo powierzchni elementów o długiej podstawie dolnej i górnej. Wystarczy więc odpowiednio rotować wykres, aby uzyskać złudzenie, że kategoria znajdująca się na dole wykresu jest większa niż ta znajdująca się na górze.

Z kolei wykresy 3D skutecznie utrudniają precyzyjne porównanie wartości poszczególnych kategorii. Jeśli przy tym stosujemy wyłącznie tzw. efekt 3D w wykresie dwuwymiarowym, problemem jest jedynie odniesienie wysokości słupka danych do wartości na osi Y. Jeśli jednak używamy trzeciego wymiaru do pokazywania danych (np. porównujemy ze sobą różne kategorie zmiennych zorganizowane wzdłuż osi Z), możemy być pewni, że czytelnik nie odczyta takiego przekazu. Poza bardzo specjalistycznymi zastosowaniami (np. wykresy powierzchniowe w fizyce czy geodezji) trzeci wymiar w wizualizacji jest po prostu zakazany.

Proponujemy również zrezygnować z wykresów skumulowanych, o ile nie służą one przedstawieniu dokładnie tego, czemu służą (czyli danych o przyroście jakiejś wielkości w stosunku do pewnego punktu początkowego; **[patrz obraz 8.22, str. 113]**). Starajmy się uważać z wykorzystywaniem koloru **[patrz podrozdział 8.2.5. „Zasada 5”, str. 118]**. Nie różnicujmy też powierzchni i objętości elementów, które chcemy ze sobą porównywać – odbiorcy naszego przekazu nie będą w stanie uchwycić tych różnic, a nasi oponenti mogą wykorzystać takie zabiegi do zakwestionowania naszej uczciwości.

### 8.2.7. Zasada 7. Opowiedz historię

Jest to najtrudniejsza do zdefiniowania, ale być może najistotniejsza ze wszystkich proponowanych tu zasad. Kluczową rolą wizualizacji jest w końcu to, aby a) pozwolić odbiorcy „zobaczyć” omawiany przez nas problem, b) wciągnąć widza w naszą narrację. W osiągnięciu obu tych celów pomaga odpowiednie zaaranżowanie wizualizacji. Do pewnego stopnia wykorzystywane przy tym środki mogą być sprzeczne z innymi omawianymi wyżej regułami. Granica, poza którą nie powinno się ich naginać, jest płynna i zależna od tego, co, do kogo i po co mówimy. W raportach dotyczących konkretnych dylematów polityki publicznej precyzja i rzetelność wizualizacji ma zapewne większe znaczenie niż np. w kampaniach profrekwencyjnych, w których usiłujemy po prostu zachęcić ludzi do jakiegoś zachowania czy zasygnalizować jakąś ogólną kwestię. Wtedy, kiedy omawiamy drastyczne dysproporcje w danych, skorzystanie z atrakcyjnego wykresu kołowego jest zapewne bardziej uprawnione, niż gdy różnice w danych nie są tak wyraziste itd. Zawsze jednak ważne jest umiejętne połączenie uczciwości w raportowaniu danych z atrakcyjną formą, uprawiającą to, że czytelnik zada sobie trud, aby odczytać nasz komunikat.

Można wymienić kilka (niesprzecznych z naszymi zasadami) sposobów na dodanie elementów historii do wizualizacji. Nawet najprostsze z tych metod bazują na założeniu, że każda dobra narracja ma jakiś punkt wyjścia, do czegoś dąży, układa fakty w pewną czytelną strukturę i kładzie na nie różny akcent. Wystarczy zatem, jeśli posortujemy nasz wykres słupkowy według wartości przedstawianej zmiennej, aby z nudnej i nieczytelnej grafiki przekształcił się on w rodzaj rankingu, w którym poszczególne jednostki rywalizują ze sobą. Już gdy zaakcentujemy istotną dla nas część wykresu odpowiednim kolorem, wyznaczymy jego punkt ciężkości. A kiedy graficznie połączymy ze sobą kilka wykresów, już utworzymy z nich opowieść.

Najbliższej klasycznej konwencji opowieści są te formy wizualizacji, które wiążą się ze wskazaniem pewnej sekwencji danych. Może to być sekwencja czasu (wykresy dynamiki jakiegoś zjawiska, animacje zmian w czasie), zdarzeń (np. wykres na osi czasu – tzw. timeline) lub przestrzeni (albo ich połączenia – jak np. słynna wizualizacja marszu armii Napoleona w kampanii rosyjskiej 1812-1813 r., uznawana przez wielu za pierwszą nowożytną infografikę<sup>44</sup>). Może to być jednak także sekwencja struktury (np. wykresy pokazujące, jaki jest proporcjonalny udział jakiejś jednostki w różnych większych całościach – por. słynna wizualizacja poziomu promieniowania<sup>45</sup>, jakie otrzymujemy z różnych źródeł). Najbardziej wyrafinowane formy graficznej narracji to coraz popularniejsze dziś rodzaje infografik – kompilacje wykresów, czyli tzw. mash-upy [patrz podrozdział 9.3. „Formaty raportów”, str. 135] lub animacje przedstawiające nie tylko dane, ale także ich wzajemne relacje.

---

44 <http://en.wikipedia.org/wiki/File:Minard.png>

45 <http://xkcd.com/radiation>

## 8.3. Narzędzia do wizualizacji danych

Skoro już znamy zasady dobrej prezentacji danych, pozostaje pytanie, jakich narzędzi możemy użyć, żeby zastosować je w praktyce? Podstawowym problemem utrudniającym udzielenie tu jakichkolwiek dobrych rad nie jest brak takich narzędzi, ale przeciwnie: ich mnogość. Biorąc pod uwagę ich dynamiczny rozwój można zakładać, że rok po opublikowaniu tego podręcznika nasze dzisiejsze narzędziowe sugestie będą już nieaktualne czy przynajmniej niekompletne. Dlatego zachęcamy wszystkich czytelników do samodzielnych poszukiwań tych narzędzi, które będą najlepiej dostosowane do ich potrzeb i możliwości. Szczególnie ciekawym źródłem może być tu np. strona organizacji od lat specjalizującej się w wizualizacji danych w kontekście działań rzeczniczych: Tactical Technology Collective ([www.tacticaltech.org](http://www.tacticaltech.org)).

W tym miejscu prezentujemy jedynie wybrane narzędzia, których sami używamy lub które wydają się nam interesującym wyborem dla organizacji: są darmowe (albo przynajmniej w pewnym zakresie bezpłatne), proste w użyciu i oferują potężne możliwości. Poniżej zamieściliśmy ich krótkie opisy – zainteresowanych odsyłamy do stron internetowych, na których można znaleźć więcej informacji, a przede wszystkim samodzielnie wypróbować polecane narzędzia.

### 8.3.1. Aplikacje dedykowane graficznej prezentacji różnych typów danych

---

#### Wykresy Google

Bardzo rozbudowany zestaw narzędzi do tworzenia wizualizacji oferuje Google. Są one dostępne dla wszystkich posiadaczy konta w tym serwisie (bezpłatnie – przy czym jednak korzystanie z aplikacji oznacza przesłanie danych do serwerów Google). Aby uzyskać odpowiedni dostęp, trzeba utworzyć arkusz kalkulacyjny (lub importować np. z Excela) na przypisanym do naszego konta dysku Google. W aktywnym arkuszu danych wybieramy następnie z menu opcję wstaw à wykres. Aplikacja sama podpowiada, jakie typy wykresów wydają się adekwatne dla naszych danych, dbając przy tym o to, aby ich wstępne formatowanie było pozbawione błędów, o których napisano wyżej. Możemy korzystać z wszelkiego rodzaju wykresów: liniowych, słupkowych, rozrzutu, histogramów, map (na poziomie krajów), Gantta oraz innych, bardziej wymyślnych. Niektóre z oferowanych typów wykresów są dość unikalne – by wspomnieć chociażby o animowanych wykresach przedstawiających w interaktywny sposób zmiany w wartości wskaźników w czasie wzorowanych na znanym projekcie Gapminder ([www.gapminder.org](http://www.gapminder.org)).

---

---

## Fusion tables

Inne ciekawe narzędzie Google to tzw. fusion tables – rodzaj tablicy, dzięki której możemy stworzyć z naszych danych np. wykres sieciowy lub interaktywną mapę (punktową) z edytowalnymi wizytówkami, możliwością filtrowania danych i wieloma innymi udogodnieniami. Potencjał fusion tables jest jednak dużo większy – umożliwiają one bowiem nie tylko wizualizację własnych danych, ale także skorzystanie z wielu danych upublicznianych za pomocą usługi (innym takim projektem jest Google Public Data Explorer, jest to jednak nie tyle aplikacja służąca do wizualizacji, co repozytorium danych).

---

## Inne usługi Google

Poza opcjami dostępnymi z poziomu dysku Google, możemy też skorzystać z różnych usług korporacji niewymagających od nas korzystania z własnych danych. Ciekawym przykładem, przydatnym zwłaszcza, jeśli kampania, którą promujemy, ma globalny charakter lub globalną wymowę, jest Google Trends ([www.google.com/trends](http://www.google.com/trends)) – wyszukiwarka umożliwiająca nam analizę haseł wpisywanych w wyszukiwarkę Google od 2004 roku. Możemy dzięki niej śledzić geografie popularności, zmiany w czasie, powiązania z innymi hasłami itp. – wszystko za pomocą kilku kliknięć. Innym ciekawym rozwiązaniem jest Google NGram Viewer (<http://books.google.com/ngrams>) – analizator działający podobnie do Google Trends, ale przeszukujący nie Internet, a zdigitalizowane przez Google książki (obecnie blisko milion pozycji).

---

## Many Eyes

To eksperymentalny projekt korporacji IBM, którego celem jest nie tylko dostarczenie użytkownikom różnych możliwości graficznej prezentacji danych, ale też stworzenie globalnej społeczności zasilającej serwis danymi (które następnie są udostępniane dla wszystkich, choć można je usunąć), tworzącej na ich podstawie wizualizacje oraz komentującej ich treść. Projekt został uruchomiony w 2007 roku przez wydział badawczy IBM i od tamtej pory znacząco się rozwinął. Oferuje uporządkowane w przyjazny sposób środowisko umożliwiające wygenerowanie niemal dowolnych rodzajów wykresów: punktowych, diagramów sieci, słupkowych, bąbelkowych, liniowych, wstęgowych, kołowych, hierarchicznych (*treemap*), chmur słów lub tagów, drzew ukazujących relacje między fragmentami testu, a także (w ograniczonym stopniu) map. Gotowe wykresy można zagnieździć na własnej stronie internetowej lub dzielić się nimi przez Facebooka lub Twittera.

Więcej: [www-958.ibm.com/software/analytics/manyeeyes](http://www-958.ibm.com/software/analytics/manyeeyes)

---

---

## Tableau Public

Projekt podobny pod pewnymi względami do przedsięwzięcia IBM rozwijany przez Tableau Software to publiczna wersja oprogramowania do analizy i wizualizacji danych przeznaczony dla środowiska biznesowego. Działa poprzez aplikację, którą trzeba ściągnąć na swój komputer. Za jej pomocą możemy utworzyć niemal dowolny wykres lub mapę (a także mozaikę złożoną z wykresów różnych typów), korzystając z niezwykle intuicyjnego interfejsu graficznego metodą „przeciągnij i upuść”. Korzystanie z aplikacji jest bardzo proste, a efekty potrafią oszałamiać – za pomocą kilku kliknięć jesteśmy w stanie osiągnąć efekty, do których wcześniej potrzebne były całe godziny pracy. Wizualizacje są nie tylko ładne, ale także interaktywne – możemy je, korzystając z filtrów, podświetlać istotne elementy danych itp. Darmowa wersja Tableau nie umożliwia jednak zapisania efektów naszej pracy na lokalnym dysku – możemy jedynie załadować gotowy wykres lub mozaikę na serwer korporacji i stamtąd zagnieździć ją na naszej stronie lub udostępnić za pomocą linków. Dla niektórych minusem może być także to, że wraz z wizualizacją udostępniamy całemu światu wszystkie dane, które są z nią powiązane. Publiczna wersja ma też limit ilości danych, choć bardzo mało uciążliwy – nie można na niej przedstawiać zestawień dla danych przekraczających milion wierszy.

Więcej: [www.tableausoftware.com/public](http://www.tableausoftware.com/public)

---

## StatSilk

To w zasadzie nie projekt, a nazwa firmy rozwijającej całą rodzinę powiązanych ze sobą programów do wizualizacji danych: StatWorld, StatTrends i StatPlanet. Pierwsza z wymienionych aplikacji służy eksploracji udostępnianych przez różne źródła (np. Bank Światowy, OECD) informacji dotyczących przemian różnych regionów świata. Dane prezentowane są w postaci mozaiki wykresów składającej się z mapy i wykresu zmian w czasie. Można dowolnie porównywać między sobą różne obszary, wybierać zakres itp. Można też podzielić się z innymi hiperłączem do określonej wizualizacji lub eksportować wykres, a także dane, na których jest oparty. StatTrends to rozbudowana wersja interaktywnej, animowanej przeglądarki danych dotyczących dynamiki zjawisk w czasie. Oferuje możliwości zbliżone do wykresu ruchomego Google. StatPlanet to z kolei aplikacja umożliwiająca interaktywną analizę danych na mapach. StatWorld jest ogólnodostępny, a rozbudowane wersje StatTrends i StatPlanet w wersjach *desktop* (aplikacji ściągniętych na dysk komputera) są dostępne nieodpłatnie dla celów osobistych oraz dla organizacji pozarządowych.

Więcej: [www.statsilk.com](http://www.statsilk.com)

---

### 8.3.2. Oprogramowanie do tworzenia różnych typów analiz/wykresów

---

#### **Wordle**

To ogólnodostępna, darmowa aplikacja służąca do przedstawiania dokumentów tekstowych w postaci chmur słów. Pozwala na stworzenie chmury nawet dla bardzo długich dokumentów, a przy tym jest niezwykle prosta w użyciu: wystarczy skopiować dowolny fragment tekstu w okno w przeglądarce i uruchomić program. W bardziej zaawansowanym trybie możemy importować do programu listę słów wraz z częstotliwością ich występowania i stworzyć z nich chmurę (pomocne w pracy z językiem polskim, w którym problemem jest deklinacja). Możemy też edytować tło chmury, kolor, czcionkę i ustawienie wyrazów itp.

Więcej: [www.wordle.net](http://www.wordle.net)

---

#### **Freeplane, lub XMind lub Mind42)**

To narzędzia służące do tworzenia mapy myśli – pomocne wtedy, kiedy chcemy ułożyć nasze dane w drzewo i pokazać ich strukturę np. podczas prezentacji. Mind42 jest mapą działającą w przeglądarce – możemy ją współdzielić i współtworzyć z innymi.

Więcej: <http://freeplane.sourceforge.net>, [www.xmind.net](http://www.xmind.net) lub <http://mind42.com>

---

#### **Batchgeo, Geocommons, Mapstory**

Są to przykłady serwisów internetowych koncentrujących się na wizualizacji danych w przestrzeni. Batchgeo umożliwia wizualizację danych o lokalizacji różnych zdarzeń lub instytucji (w wersji darmowej do 250 zdarzeń). Aby jej użyć, wystarczy wkleić w okno na głównej stronie tekst z danymi adresowymi lokalizacji, która nas interesuje. Aplikacja sama „geokoduje” adresy (przypisuje im atrybuty przestrzenne) i prezentuje na mapie w postaci klasycznych pinezek wraz z informacjami z naszych danych dostępnymi w dymkach. Geocommons to serwis bazujący na oprogramowaniu GeoIQ, gromadzący mapy przygotowane przez użytkowników i umożliwiający tworzenie własnych map. Oferuje wiele interesujących funkcji, w tym analizę danych geograficznych w czasie, współdzielenie i zagnieźdżanie wizualizacji, dostęp do otwartych danych różnego typu, API dla deweloperów aplikacji internetowych itp. Mapstory to specyficzna aplikacja dedykowana tworzeniu i współdzieleniu animowanych map ukazujących zmiany w czasie.

Więcej: <http://batchgeo.com>, <http://geocommons.com> lub <http://mapstory.org>

---

---

### **Gephi (lub Pajek, Netlogo)**

To bardzo zaawansowane narzędzia umożliwiające tworzenie interaktywnych wizualizacji sieci (np. powiązań między członkami jakiejś grupy, stronami internetowymi itp.). Gephi jest bardzo proste w użyciu (pod warunkiem, że opanowaliśmy podstawy analizy sieciowej) i umożliwia tworzenie wysokiej jakości wykresów sieci oraz wyliczanie na ich podstawie istotnych statystyk.

Więcej: <https://gephi.org>, <http://vlado.fmf.uni-lj.si/pub/networks/pajek>  
lub <http://ccl.northwestern.edu/netlogo>

---

### **TimelineJS lub Dipity**

Jest to oprogramowanie do tworzenia wizualizacji typu oś czasu – a więc przewijanych diagramów ukazujących kalendarium wydarzeń w czasie. Obie aplikacje umożliwiają tworzenie i współdzielenie atrakcyjnych wizualnie tego rodzaju grafik, do których można dodawać obrazy, zdjęcia lub materiały wideo, opisy itp.

Więcej: <http://timeline.verite.co> lub [www.dipity.com](http://www.dipity.com)

---

### **Ushahidi**

Jest to wolne oprogramowanie typu *crowdmapping* – przeznaczone do nanoszenia na mapy informacji przez wielu użytkowników. Rozwijane jest przez organizację non profit specjalizującą się w tworzeniu wolnego oprogramowania służącego do szybkiego przetwarzania dużych ilości informacji. Pierwsza wersja Ushahidi powstała jako platforma do raportowania przypadków przemocy po wyborach w Kenii w 2008 roku (stąd nazwa, która w suahili oznacza „świadeństwo”). Wykorzystywano ją także w trakcie akcji ratunkowej po trzęsieniu ziemi na Haiti, a także w Polsce w czasie powodzi w 2010 r. Umożliwia szybkie dodawanie raportów o różnych zdarzeniach (podzielonych na kategorie) za pomocą komputera (mailem, przez Twittera, witrynę internetową) lub telefonu. Dane te są wyświetlane na mapie i wykresie ukazującym natężenie raportów w czasie. Platforma jest całkowicie darmowa, ale wymaga zainstalowania na własnym serwerze. W nagłych przypadkach lub do szybkich działań można jednak skorzystać z usługi Crowdfunder, która jest uproszczoną wersją aplikacji Ushahidi możliwą do uruchomienia w kilka minut.

Więcej: [www.ushahidi.com](http://www.ushahidi.com)

---

---

## **QuantumGIS lub Mapnik**

Są to zaawansowane aplikacje typu GIS, służące nie tyle do samej prezentacji danych na mapach, ale do zaawansowanej analizy informacji przestrzennej. To otwarte, dostępne za darmo odpowiedniki bardzo kosztownych programów typu ArcGIS czy Mapinfo. Polecamy je przede wszystkim użytkownikom mającym duże doświadczenie w analizach przestrzennych i posiadającym niezbędne do nich dane.

Więcej: [www.qgis.org](http://www.qgis.org) lub <http://mapnik.org>

---

### **8.3.3. Programy do robienia infografik**

---

#### **Infogr.am**

Coraz częściej chcemy komunikować się ze światem nie za pomocą skomplikowanych wykresów, ale poprzez bardzo proste, a zarazem ładne infografiki. W takich sytuacjach nieocenioną pomocą jest Infogr.am – serwis umożliwiający użytkownikom opracowywanie własnych infografik na podstawie gotowych, edytowalnych szablonów (niestety nie jest ich zbyt wiele, ale do podstawowych zastosowań z pewnością wystarczą). Korzystając z serwisu, możemy utworzyć własną tablicę złożoną z wielu wysokiej próby grafik i wysłać ją w świat lub zagnieździć na własnej stronie. Prostota interfejsu sprawia, że autorem infografiki może zostać praktycznie każdy, kto ma pomysł na własny przekaz i jakiegokolwiek dane na podorzędziu.

Więcej: <http://infogr.am>

---

#### **Piktochart**

Jest to aplikacja bliźniacza w stosunku do Infogr.am. Oferuje podobne możliwości, a do tego umożliwia eksport wysokiej rozdzielczości obrazów z infografikami. Udostępnia też znacznie szerszą paletę szablonów infografik i wykresów. Jest nieodpłatna tylko w podstawowym zakresie (do 10 obrazów, 7 szablonów, nieusuwalne logo Piktochart na infografice). Jednak dla organizacji pozarządowych przewidziano możliwość wykupienia abonamentu za ok. 40 dolarów rocznie.

Więcej: <http://piktochart.com>

---



## 9. Raporty rzecznicze

Zebranie danych, przeanalizowanie ich i graficzne przedstawienie to jeszcze nie koniec pracy. Musimy je jeszcze w jakiś sposób przekazać i ubrać w formę, która będzie spójna i czytelna dla odbiorcy. Tak przygotowane dane powinny stać się ważną częścią raportu rzeczniczego, który często stanowi zwieńczenie prac nad przygotowaniem i sformułowaniem naszych postulatów.

Pod pojęciem „raport” kryje się mnóstwo najrozmaitszych form, różniących się od siebie charakterem, strukturą, wielkością, językiem, jakim są napisane, grupami odbiorców, nie mówiąc już o rozwiązaniach graficznych. Niniejszy rozdział nie jest próbą stworzenia systematycznego i wyczerpującego opisu różnych form, jakie mogą przybierać raporty rzecznicze. Chcemy jednak wskazać na kilka wątków, które mogą mieć znaczenie przy tworzeniu raportów.

### 9.1. Rodzaje raportów

Nie istnieje żadna powszechnie uznana klasyfikacja raportów, można jednak wyodrębnić kilka ich rodzajów.

Najbardziej klasyczną formą publikacji przedstawiającej stanowisko rzecznicze jest tzw. **policy paper** – pojęcie to nie ma swojego odpowiednika w języku polskim. Jest to względnie zwarte, kilkunasto- lub kilkudziesięciostronicowe opracowanie, w którym przedstawia się określony problem w świetle danych oraz na tej podstawie proponuje się pewne zmiany w prawie czy szerzej w politykach publicznych. Z reguły jest szczególną formą „propozycji” kierowaną do polityków lub przedstawicieli administracji.

---

#### „Samorządność uczniowska w systemie edukacyjnym” autorstwa

Jędrzeja Witkowskiego Instytut Spraw Publicznych, Warszawa 2013.

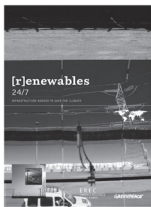
Jest to klasyczny **policy paper**, będący efektem prac „Koalicji na rzecz samorządów uczniowskich”. Został wydany przez Instytut Spraw Publicznych – fundację specjalizującą się w tego typu opracowaniach. Autor przedstawia najpierw wagę problemu („znaczenie samorządności uczniowskiej”), następnie sytuację w Polsce („stan samorządności uczniowskiej w Polsce”), po czym przywołuje zapisy obowiązujących ustaw i w odniesieniu do nich formułuje postulaty zmian (w postaci konkretnych zapisów).

Dokument dostępny jest na stronie: [www.ceo.org.pl/sites/default/files/news-files/samorznosc\\_uczniowska\\_w\\_systemie\\_educacyjnym.pdf](http://www.ceo.org.pl/sites/default/files/news-files/samorznosc_uczniowska_w_systemie_educacyjnym.pdf)

---



Innym typem jest **raport ekspercki** – przygotowany na ogół przez ekspertów i dla ekspertów. Dotyczy z reguły szczegółowych zagadnień w obrębie pewnego problemu, które zrozumiałe są w pełni dla osób naprawdę znających się na jakimś temacie. Najczęściej tego typu opracowania pojawiają się w przypadku twardych dziedzin, takich jak np. energetyka czy finanse publiczne.



### **„[R]enewables 24/7 – Infrastructure needed to save the climate”**

Greenpeace International 2009.

Raport Greenpeace o odnawialnych źródłach energii jest tego typowym przykładem raportu eksperckiego. Publikacja jest gęsto ilustrowana wykresami, pełna finansowych wyliczeń i technicznych terminów – jest przez to mało przystępna dla laika, nawet jeśli ten interesuje się tematem, ale stanowi zapewne bardzo rzeczowy głos w dyskusji między ekspertami.

Raport dostępny jest na stronie: <http://www.greenpeace.org/poland/Global/poland/report/2010/2/onawialne-r-d-a-energii-24-7.pdf>

Często za raport rzeczniczy posłużyć może **raport z badania**. Określone badanie dedykowane kwestiom istotnym dla zakresu działań rzeczniczych staje się podstawą i punktem wyjścia. To na jego wynikach koncentruje się opis – ale przy tym rozbudowujemy wnioski i staramy się przełożyć je na rekomendacje dotyczące działania (najczęściej działania władz).



### **„Przyjazny Urząd? Dostęp do informacji publicznej na wniosek w gminach i miastach województwa śląskiego”**

Stowarzyszenie Wzajemnej Pomocy BONA FIDES, Katowice 2012.


Stowarzyszenie BONA FIDES przeprowadziło w 2012 roku badania we wszystkich gminach województwa śląskiego, sprawdzając sposób wywiązywania się urzędów z obowiązku udostępniania informacji publicznej. Publikacja przedstawia wyniki badania, ale zawiera również rozdział „Propozycje rozwiązań”, w którym autorzy przedstawiają krótko i bardzo konkretnie, co ich zdaniem powinny zrobić urzędy, by poprawić swoje funkcjonowanie w tym obszarze.

Funkcję rzeczniczą mogą pełnić pośrednio również różnego rodzaju **podręczniki i poradniki**, zwłaszcza te skierowane do administracji. Ich celem jest przekonać do określonych działań poprzez wręczenie gotowych narzędzi. Urzędnik lub polityk, który miałby pójść postulowaną

przez organizację drogą ma więc o jeden ważny problem mniej, gdy od razu wie, w jaki sposób działać. Przykład tego typu podręcznika przywołujemy poniżej, przy okazji opisywania struktury raportów rzeczniczych („Dobry start” Fundacji im. J. A. Komeńskiego).

Podobną rolę pełnią **opisy dobrych praktyk**. Z reguły mają one zachęcić do naśladowania opisywanych procesów – pokazują, jak to robić i stanowią dowód, że proponowane rozwiązanie rzeczywiście działa. Trzeba jednak pamiętać, że jeśli sami nie uczestniczyliśmy w opisywanych zdarzeniach, przygotowanie takiej publikacji wymaga gruntownego zapoznania się z ich przebiegiem.

---



**Baza dobrych praktyk partycypacji projektu „Decydujmy razem”**  
Fundacja Inicjatyw Społeczno-Ekonomicznych

Udział obywateli w tworzeniu polityk publicznych  
INICJATYWY SPOŁECZNO-EKONOMICZNE  
FISE

W bazie zebrano opisy różnych procesów z zakresu partycypacji obywatelskiej, z Polski i zagranicy. Wszystkie przykłady zostały opisane na podstawie wizyt studyjnych, w formie, która ma sprzyjać skorzystaniu z bazy przez urzędników: z rozpisаныmi poszczególnymi etapami procesu, z wyliczonym budżetem, z przedstawionymi korzyściami oraz barierami, na jakie się można natknąć w trakcie realizacji podobnych działań.

Na podstawie bazy powstały dwie publikacje: „Udział obywateli w tworzeniu polityk publicznych – wybór dobrych praktyk zagranicznych” (2010) oraz „Elementy partycypacji” (2012). Baza dostępna jest na stronie: <http://dobrepraktyki.decdujemyrazem.pl>

---

## 9.2. Struktura raportu

Zależnie od rodzaju raportu, który przygotowujemy, różnie może wyglądać jego struktura. Są jednak pewne „żelazne punkty”, które powinny znaleźć się w opracowaniu przedstawiającym nasze stanowisko w oparciu o dane. Każdy z tych elementów spełnia ważną funkcję, która ma przysłużyć się naszej działalności rzeczniczej.

**Podstawowy komunikat.** W miejscu, na które łatwo trafić – na okładce lub na którejś z pierwszych stron – powinniśmy zwięźle napisać to, co uważamy za istotę naszego komunikatu. Powinno to być jedno, maksymalnie kilka zdań. Często wydaje nam się, że tak skrótowo nie uda ująć się tego, co chcemy przekazać, bo problem jest złożony. Jednak pamiętajmy, że przede

wszystkim powinniśmy zachęcić kogoś, kto weźmie do ręki naszą publikację, żeby zechciał ją chociaż przekartkować. Nasi potencjalni odbiorcy mogą mieć do czynienia z dziesiątkami raportów – musimy więc sprawić, żeby zainteresowali się akurat naszym. Sam **tytuł raportu** może zadecydować o tym, czy ktoś weźmie go do ręki. Spróbujmy więc wymyślić taki, który z jednej strony przyciągnie potencjalnego czytelnika, a z drugiej uchwyci istotę naszego przekazu.

**Problem.** Bardzo ważne jest wyraźne określenie problemu, którym się zajmujemy. Może to się wydawać oczywiste – ale często zdarza się, że nie tłumaczymy czegoś, bo uważamy to za zrozumiałe samo przez się. Musimy jednak pamiętać, że dla odbiorców to wcale takie nie musi być. Powinniśmy wyjaśnić, dlaczego zajmujemy się daną kwestią i dlaczego uważamy ją za ważną. Dzięki temu ktoś, kto weźmie do ręki naszą publikację, będzie mógł lepiej zrozumieć, o co nam chodzi.

**Rekomendacje/postulaty.** Naszą publikacją chcemy przekonać do podjęcia lub zaniechania określonych działań. Często jednak – zwłaszcza wtedy, kiedy skupiamy się na analizie danych – możemy uznać, że z naszego opisu sytuacji w sposób oczywisty wynika, jakie są jej praktyczne konsekwencje. Podobnie jak w przypadku ogólnego problemu, może to być pochopne założenie. Powinniśmy w osobnej części opisać, najlepiej w zwięzłych, precyzyjnych punktach, jakie są nasze postulaty i jaki jest proponowany sposób ich realizacji.

**Argumentacja.** Dane nigdy nie mówią też same za siebie – zawsze potrzebują omówienia, które umieści je w odpowiednim kontekście. Dlatego pamiętajmy, że nie wystarczy zarzucić czytelnika danymi – czasami lepiej przedstawić ich mniej, ale precyzyjnie wybranych. Powinniśmy pokazać, dlaczego przytoczone informacje są ważne, jak odnoszą się do tematu, co z nich wynika i jaki jest związek pomiędzy nimi a naszymi postulatami. Krótko mówiąc – należy zamienić dane w spójne argumenty.

**Dookreślenie odbiorców.** Z reguły naszych raportów nie kierujemy do wszystkich, lecz do określonych grup, np. polityków, przedstawicieli innych organizacji – a czasem jeszcze konkretniej, np. tylko urzędników z określonych wydziałów w urzędach gminy. Dobrze jest to wyraźnie napisać na początku publikacji. Jeśli urzędnik będzie wiedział, że raport jest skierowany specjalnie do niego, może chętniej zainteresować się jego treścią.

**Informacje o organizacji.** W rzecznictwie bardzo ważne jest, kto stoi za danymi działaniami. Dlatego nie powinniśmy być anonimowi, lecz opisać naszą organizację, żeby odbiorcy wiązali

z nią treść publikacji. Najczęściej kilku- lub kilkunastozdaniowy opis znajduje się na którejś z ostatnich stron.

**Informacje o autorach.** Z tego samego powodu powinniśmy przedstawić najważniejsze informacje o autorach danej publikacji. W rzecznictwie nie możemy być anonimowi. Z jednej strony, raport sygnowany rozpoznawalnym nazwiskiem lub charakterystyczną marką ma większą rangę; z drugiej, kolejne działania powiększają reputację autorów i/lub organizacji, co daje większe możliwości na przyszłość.

**Sponsorzy/źródła finansowania.** Bardzo ważna jest przejrzystość i ujawnienie, skąd pochodzą pieniądze na działalność rzeczniczą, a w szczególności na przygotowanie i promocję samego raportu. Być może jest to szczególnie ważne właśnie wtedy, kiedy sponsorzy mogą być w jakimś stopniu niewygodni. Jeśli sami nie ujawnimy takiej informacji, zawsze ktoś się jej doszuka i zrobi to za nas. Natomiast jeśli otwarcie przedstawimy naszych sponsorów, możemy być spokojni – jedyne, czego możemy się obawiać, to merytoryczna debata na temat tego, czy źródła finansowania nie wpływają na charakter naszej działalności. Nikt nie może nam jednak zarzucić, że coś ukrywamy.

**Metodologia.** Kiedy korzystamy z danych – a do tego staramy się przekonywać w niniejszym podręczniku – kluczowe jest podawanie szczegółów dotyczących wykorzystanych źródeł. Tylko wtedy możemy na ich temat sensownie dyskutować. Zależnie od charakteru publikacji, część opisująca metodologię zbierania danych może być bardziej lub mniej rozbudowana. W przypadku raportu z badań będzie szczególnie istotna. W innych przypadkach może być wskazane ograniczyć ją na tyle, żeby nie zdominowała naszego przekazu. Jednak możemy wtedy np. odesłać do strony internetowej, na której umieścimy bardziej szczegółowe informacje na temat wykorzystanych przez nas danych.

Proste połączenie powyżej wymienionych części oczywiście nie wystarczy, żeby stworzyć ciekawy i przekonujący raport rzeczniczy. Nie ma tu jednej doskonałej recepty. Ostateczna struktura każdorazowo zależy od specyfiki danego problemu i celu, który chcemy zrealizować.

Poniżej przedstawiamy kilka ciekawych przykładów tego, jak może być skonstruowany interesujący raport.



## **„Edukacja małych dzieci – standardy, bariery, szanse”, pod redakcją Anny Gizy,**

**Fundacja Rozwoju Dzieci im. J. A. Komeńskiego, Warszawa 2010.**

Raport Fundacji Komeńskiego, promujący dobre rozwiązania w zakresie edukacji małych dzieci, jest podzielony na dwie osobne części. Publikacja opiera się na rozbudowanej analizie danych – jednak jest ona zupełnie oddzielona od części,

w której prezentowane są rekomendowane rozwiązania.

Pierwsza część to narracja rzecznicza, w której żadne szczegółowe dane nie są przywoływane. Na początku zaprezentowane są „zasady dobrej edukacji”, później opisano bariery w edukacji małych dzieci, a następnie zaproponowano, jak można te przeszkody przezwyciężyć. Gdyby raport ograniczył się do tego, można by zarzucić, że zawarto w nim wiele bezpodstawnych twierdzeń. Jednak integralną i znacznie dłuższą – choć osobną – częścią publikacji jest „raport z kwerend”. Tam znajdują się wszystkie szczegółowe dane, na których opierają się stwierdzenia w części rzeczniczej.

Tego typu oddzielenie części rzeczniczej i części badawczej raportu może być ciekawym rozwiązaniem, jednak nie zawsze najlepszym. Czasami, zwłaszcza wtedy, kiedy spór toczy się o same dane i ich znaczenie, niezbędne jest podawanie ich na bieżąco – bezpośrednio w odniesieniu do stwierdzeń, które na ich podstawie formułujemy.

Raport dostępny jest na stronie: [www.frd.org.pl/repository/raporty/Raport\\_o\\_barierach.pdf](http://www.frd.org.pl/repository/raporty/Raport_o_barierach.pdf)



## **„Dobry start. Jak wprowadzać alternatywne formy wczesnej edukacji?”**

**Fundacja Rozwoju Dzieci im. J. A. Komeńskiego, Warszawa 2008.**

Kolejna interesująca publikacja Fundacji Komeńskiego ma formę podręcznika dla przedstawicieli gminnych samorządów, poświęconego rozwijaniu alternatywnych form edukacji przedszkolnej. Jego rzeczniczy charakter jest bardzo widoczny: podręcznik ma na celu popularyzację tychże alternatywnych form wśród władz samorządowych. Jego strukturę można uznać za modelową.

W pierwszym rozdziale, na podstawie licznych danych, przedstawiono znaczenie edukacji przedszkolnej, mające wskazać na wagę całego problemu. Następnie opisane są alternatywne formy edukacji przedszkolnej, jako propozycja Fundacji na rozwiązanie problemu, którym jest niskie uprzedzszkolnienie. Autorki niejako podają władzom samorządowym gotowe rozwiązanie na tacy: zakładanie ośrodków przedszkolnych jest opisane krok po kroku, a na końcu podręcznika podano również szczegóły organizacyjne i finansowe takiego przedsięwzięcia. Wreszcie, po pokazaniu, dlaczego warto to robić i jak działać, przytoczone są dane z ewaluacji

już istniejących ośrodków przedszkolnych, które pokazują, że takie rozwiązanie się sprawdza. W taki sposób autorki podręcznika dostarczyły odbiorcom – władzom samorządowym – pełnej wiedzy potrzebnej do podjęcia postulowanych przez Fundację działań.

Raport dostępny jest na stronie: [www.frd.org.pl/repository/documents/dobry\\_start.pdf](http://www.frd.org.pl/repository/documents/dobry_start.pdf)

---



### **„Public Spending for Public Benefit”**

New Economics Foundation, London 2005.

Raport NEF, brytyjskiej organizacji zajmującej się promowaniem alternatywnych modeli rozwoju, jest dla nas interesujący ze względu na pierwsze kilka stron.

Po pierwsze, zawiera podsumowanie, tzw. *executive summary* (dość częsty element różnych raportów), w którym na dwóch stronach wypunktowano najważniejsze treści z każdego rozdziału. Podobnie na początku każdego rozdziału znajduje się ramka z kluczowymi punktami, dzięki czemu nie trzeba się zastanawiać, które treści autorzy uważają za najistotniejsze.

Jednak to, co najciekawsze w tym raporcie, to tabelka, znajdująca się jeszcze przed podsumowaniem, zatytułowana „Co zyskam dzięki temu raportowi?”. W tabelce przedstawiono, dlaczego publikacja może być interesująca dla czytelnika, jeśli jest on przedstawicielem sektora publicznego, biznesu lub lokalnej społeczności. Każdą z tych trzech kategorii podzielono na podkategorie (np. w sektorze prywatnym wyróżniono „mały biznes”, „duży biznes” i „przedsiębiorstwa społeczne”). Dla każdej kategorii i podkategorii wypisano zwięzłe hasło, mające przekonać czytelnika, że powinien zapoznać się z tym raportem.

Przykład NEF jest ciekawy, ponieważ z reguły myślimy o swoich działaniach z punktu widzenia przyświecających im wartości. Uważamy, że jakiś raport powinien być przeczytany i wykorzystany np. przez polityków po prostu dlatego, że problem, który w nim opisujemy, jest istotny społecznie. NEF odwraca trochę tę logikę, a zachęcając do zapoznania się z raportem, wykorzystuje język korzyści.

Raport dostępny jest na stronie: [www.neweconomics.org/page/-/files/Public\\_Spending\\_for\\_Public\\_Benefit.pdf](http://www.neweconomics.org/page/-/files/Public_Spending_for_Public_Benefit.pdf)

---

## **9.3. Formaty raportów**

Wraz z rozwojem nowych technologii i większą dostępnością różnych narzędzi np. do prostej obróbki graficznej pojawia się coraz więcej różnych form, w których raporty są publikowane. Większe organizacje, mogące sobie pozwolić na opłacenie grafika czy informatyka, często

odchodzą od tradycyjnej formy papierowego raportu, zawierającego sam tekst lub ewentualnie wykresy, diagramy itp.

Warto zauważyć, że takie możliwości są niedostępne dla licznych organizacji zajmujących się rzecznictwem. Dla wielu z nich rzecznictwo jest dodatkiem do głównej działalności (polegającej np. na świadczeniu usług) i nie mają na nie osobnych funduszy, lecz opierają się na społecznej pracy członków lub pracowników zajmujących się w ramach swoich obowiązków innymi rzeczami. W takiej sytuacji samo zaplanowanie działań rzeczniczych i przygotowanie najprostszego raportu jest już dużym wysiłkiem.

Przywiązywanie dużej uwagi do strony graficznej raportu niesie za sobą też pewne ryzyko. Wydrukowanie go w pełnym kolorze, na papierze kredowym i z ekstrawaganckim projektem graficznym może budzić pytania, skąd organizacja ma tyle pieniędzy, żeby móc sobie pozwalać na takie „zbytki” i czy na pewno właściwie te fundusze wydaje. Może to stawiać nas w złej pozycji w dyskusji z odbiorcami raportu. Zależy to oczywiście od tego, do kogo kierujemy swój przekaz. Taki „zbytkowny” raport będzie na pewno odebrany gorzej, jeśli przygotowuje go lokalna organizacja dla władz gminy, niż w sytuacji, kiedy ogólnopolska fundacja zwraca się do polityków na szczeblu rządowym lub parlamentarnym.

Poza tym, że raporty drukowane mają coraz częściej zaawansowane graficznie formy, znaczenia nabierają również wersje elektroniczne. Umieszczanie tych publikacji na stronie internetowej w postaci pliku PDF jest już w zasadzie powszechnym standardem. Pojawiają się jednak również inne formy internetowej prezentacji publikacji, takie jak **książki internetowe**, imitujące przewracanie kartek w wersji papierowej, lub raport otwarty.

Raportem otwartym został nazwany przez autorów raport „Młodzi i Media” Szkoły Wyższej Psychologii Społecznej (<http://www.mim.swps.pl>). Jest on umieszczony w formie tekstowej na samej stronie internetowej, w podziale na poszczególne rozdziały. Przy każdym akapicie jest miejsce na to, żeby zostawić swój **komentarz**. Założeniem autorów było, żeby raport był uzupełniany przez czytelników (oraz samych twórców). Jak na razie komentarzy na takich stronach jest niewiele, ale warto wiedzieć o tego typu pomysłach. „Młodzi i Media” to nie był akurat raport rzeczniczy – jednak w podobny sposób został umieszczony w Internecie „Raport o stanie samorządności terytorialnej w Polsce” pod redakcją prof. Jerzego Hausnera (<http://stansamorządu.pl>).



Internet odgrywa coraz większą rolę również dlatego, że to tam umieszczane są rozmaite elementy towarzyszące raportom lub wręcz w niektórych sytuacjach zastępujące je. Jednym z takich elementów są coraz modniejsze infografiki – przekładające długi raport na kilka ładnie wyglądających obrazków, zawierających liczby i wykresy. Kilka organizacji wykorzystuje też tzw. **mash-upy, czyli infografiki animowane**. Tego typu graficzne rozwiązanie zostało zastosowane przy prezentacji wyników z badań przeprowadzonych przez organizację, m.in. „Po co Polakom biblioteki?” ([www.biblioteki.org/maszap](http://www.biblioteki.org/maszap)) czy „Obiegi kultury” (<http://obiegi-kultury.centrumcyfrowe.pl/mashup>).

Takie produkty towarzyszące raportom mogą przynieść bardzo pozytywne skutki, ale niosą też z sobą zagrożenia.

Z pewnością, ze względu na atrakcyjną formę, mogą łatwiej zainteresować potencjalnego odbiorcę niż suchy tekst, zwłaszcza że czytanie raportu wymaga czasu, a treść infografiki można przyswoić w ciągu kilkunastu sekund. Jednak właśnie w tym tkwi zarazem zagrożenie. Korzystając z takich sposobów prezentacji jak infografika, ryzykujemy, że dane będą jawić się przez to jako bardziej powierzchowne. Te same czynniki, które mogą zwiększać atrakcyjność przekazu, często zmniejszają jego precyzję (np. kolorowy wykres, bardziej przyciągający uwagę, mniej czytelnie przedstawia dane niż wykres czarno-biały). Ryzykujemy więc tym samym, że debata stanie się mniej merytoryczna – a cel wykorzystania danych jest przecież wręcz przeciwny.

Istnieje jeszcze jeden sposób, który wykorzystuje się, żeby spróbować dotrzeć do tych odbiorców, którzy raczej niechętnie zaborą się za lekturę długiej publikacji. Czasami powstają dwie wersje raportu: pełna oraz skrócona. W pełnej wersji zawieramy wszystko to, co uważamy za niezbędne dla potrzeb merytorycznej dyskusji – całość danych i szczegółową argumentację uzasadniającą nasze stanowisko. W wersji skróconej za to kładzie się nacisk na najważniejsze wnioski i postulaty. Każdy, kto chce się zapoznać z tokiem myślenia stojącym za nimi, może sięgnąć do wersji pełnej, gdzie znajduje się również np. część metodologiczna raportu. W niektórych przypadkach dla najbardziej dociekliwych czytelników udostępnia się pełne dane, na podstawie których powstał raport. Tak postąpili autorzy „Obiegów kultury”, którzy zamieścili na stronie projektu pakiet surowych danych z przeprowadzonych badań. Dzięki temu można je samodzielnie poddać analizie statystycznej i skonfrontować uzyskane wyniki z tymi, które przedstawiono w raporcie.

## 9.4. Rola raportu

Na koniec pozostaje ważne pytanie o to, jaka ma być rola przygotowanego przez nas raportu w działalności rzeczniczej. Celem niniejszego podręcznika było zwrócenie uwagi na znaczenie, jakie ma wykorzystanie danych w debacie publicznej. Raport jest formą prezentacji zebranych danych i opartego na nich stanowiska. Jednak z całą pewnością nie jest ostatnim etapem działań rzeczniczych – wręcz przeciwnie, najczęściej jest punktem wyjścia do właściwego rzecznictwa.

Naiwnością jest liczenie na to, że przygotujemy raport, wyślemy go do odpowiedniego posła lub radnego, który go przeczyta, zamieni nasze postulaty na projekt ustawy/uchwały i wprowadzi w życie. Zresztą większość organizacji zajmujących się rzecznictwem prawdopodobnie nie życzyłoby sobie takiego scenariusza. Dane, które zbieramy i prezentujemy w postaci raportu, mają stanowić podstawę do wspólnej rozmowy – włączającej różne zainteresowane strony – a nie stanowić ostateczne rozstrzygnięcie problemu. To ta dyskusja, wspólne uzgadnianie rozwiązań, a potem wdrażanie ich i monitorowanie – stanowią o istocie rzecznictwa. W podejściu, które prezentujemy, dane są niezbędną bazą, znajdującą się jednak bardziej w tle niż w centrum debaty.

Co więcej, dane nie tylko nie kończą rozmowy, ale często jej nawet nie rozpoczynają. Sama treść raportu nie przekona nikogo do zainteresowania się tematem – bo najpierw trzeba ten raport przeczytać. Dlatego staje się on naprawdę przydatny wtedy, kiedy odbiorca, do którego kierujemy nasz przekaz, jest już nastawiony, żeby na dany temat dyskutować. Do zainteresowania naszym przekazem musimy więc wykorzystać inne sposoby.

Oba te etapy działań – po pierwsze, formy promocji mające wzbudzić zainteresowanie przygotowanym przez nas stanowiskiem, a po drugie, prace nad przygotowaniem i wdrożeniem ostatecznych rozwiązań – to jednak tematy na zupełnie inny podręcznik. Mamy nadzieję, że również on wkrótce powstanie.

## 10. Bibliografia

### ROZDZIAŁ 2

- M. Borowska, M. Branka, J. Wcisło, *Docenić bibliotekę. Jak skutecznie prowadzić rzecznictwo*, Fundacja Rozwoju Społeczeństwa Informacyjnego, Warszawa 2012.
- Legislacja w samorządzie terytorialnym. Monitorowanie prawa miejscowego przez przedstawicieli społeczności lokalnych*, Fundacja Rozwoju Demokracji Lokalnej, Warszawa 2010.
- A. Rajner, J. Wygnańska (red.), *Rzecznictwo: jak skutecznie zabiegać o realizację praw osób z autyzmem i innymi całościowymi zaburzeniami rozwojowymi*, Fundacja SYNAPSIS, Warszawa 2004.
- R. Skrzypiec, „Współpraca z otoczeniem”, w: *Wiedza i doświadczenie*, moduł III: *Współpraca z otoczeniem*, Akademia Rozwoju Filantropii w Polsce, Warszawa 2006.
- „Trzeci sektor” 6/2006: „Rzecznictwo i lobbying organizacji pozarządowych: Doświadczenia rzecznicze polskich organizacji”.

### ROZDZIAŁ 5

#### Publikacje

- E. Babbie, *Badania społeczne w praktyce*, PWN, Warszawa 2007.
- Dostęp do danych publicznych w Polsce. Raport z monitoringu Stowarzyszenia Klon/Jawor*, Stowarzyszenie Klon/Jawor, Warszawa 2013.
- Ewaluacja. Jak to się robi?*, Pracownia Badań i Innowacji Społecznych „Stocznia”, Warszawa 2013.
- J. Gray, L. Bounegru, L. Chambers (red.), *The Data Journalism Handbook. How Journalists Can Use Data to Improve the News*, O’Reilly Media, Sebastopol 2012.
- S. Mandes, „Metody jakościowe w ewaluacji”, w: K. Olejniczak, M. Kozak, B. Ledzion (red.), *Teoria i praktyka ewaluacji interwencji publicznych*, Wydawnictwa Akademickie i Profesjonalne, Warszawa 2008.
- Ł. Widła, „Zastosowanie metod ilościowych w ewaluacji”, w: K. Olejniczak, M. Kozak, B. Ledzion (red.), *Teoria i praktyka ewaluacji interwencji publicznych*, Wydawnictwa Akademickie i Profesjonalne, Warszawa 2008.

#### Strony internetowe

- Archiwum Danych Społecznych: <http://www.ads.org.pl/>
- Bank Światowy: <http://data.worldbank.org/>
- Baza raportów CBOS: <http://www.cbos.pl/PL/publikacje/raporty.php>
- Diagnoza społeczna: <http://www.diagnoza.com/>
- Eurostat: <http://ec.europa.eu/eurostat>
- Główny Urząd Statystyczny: <http://www.stat.gov.pl/>

Mam Prawo Wiedzieć: <http://www.mamprawowiedziec.pl/>

Moja Polis: <http://www.mojapolis.pl/>

Nasza Kasa: <http://www.naszakasa.org.pl/>

OECD: <http://stats.oecd.org/> oraz <http://www.oecd.org/statistics/>

Pozarządowe Centrum Dostępu do Informacji Publicznej: <http://informacjapubliczna.org.pl>

Sejmometr: <http://sejmometr.pl/>

## **ROZDZIAŁY 6, 7 i 8**

### **Publikacje**

*Elektroniczny Podręcznik Statystyki PL*, Statsoft, Kraków 2006.

H. Kassyk-Rokicka, *Statystyka nie jest trudna. Mierniki statystyczne*, PWE, Warszawa, wiele wydań.

J. Kennedy, *Principles of Information Visualization Tutorial – Part 1. Design Principles*, Institute for Informatics & Digital Innovation.

N. Yau, *Visualise This! The FlowingData Guide to Design, Visualization, and Statistics*, Wiley Publishing, Indianapolis 2011.

### **Strony internetowe**

Drawing by Numbers: <http://drawingbynumbers.org/>

Na Straży Sondaży: <http://nastrazysondazy.uw.edu.pl/>

School of Data: <http://schoolofdata.org/>

Blog Smarter Poland: <http://smarterpoland.pl/>

