

UNIwersytet Szczeciński
ROZPRAWY I STUDIA T. (MCCCLXXI) 1297

Ewa Krok

Wizualizacja danych, informacji i wiedzy

Szczecin 2023

Rada Wydawnicza

Barbara Braid, Anna Cedro, Urszula Chęcińska, Rafał Klóska
Maciej Kowalewski, Ewa Mazur-Wierzbicka, Jarosław Nadobnik
Grzegorz Wejman, Renata Ziemińska, Magdalena Zioło
Andrzej Skrendo – przewodniczący Rady Wydawniczej
Elżbieta Zarzycka – dyrektor Wydawnictwa Naukowego

Recenzent

dr hab. Helena Dudycz, prof. UEW

Redakcja językowa

Elżbieta Blicharska

Korekta

Paulina Iwan

Skład

Joanna Dubois-Mosora

Projekt okładki

Joanna Dubois-Mosora

Ewa Krok ORCID 0000-0003-1512-6901



Wersja elektroniczna publikacji dostępna na licencji CC BY-SA 4.0
po 12 miesiącach od daty wprowadzenia do obrotu

© Copyright by Uniwersytet Szczeciński, Szczecin 2023

DOI 10.18276/978-83-7972-673-8

ISBN 978-83-7972-673-8 (online)

ISBN 978-83-7972-683-7 (print)

ISSN 0860-2751

WYDAWNICTWO NAUKOWE UNIWERSYTETU SZCZECIŃSKIEGO

Wydanie I | Ark. wyd. 8,0 | Ark. druk. 7,8 | Format A5.

SPIS TREŚCI

Wprowadzenie	5
1. Elementy tworzenia użytecznych wizualizacji.....	7
1.1. Definicje wizualizacji	7
1.2. Cel wizualizacji	10
1.3. Treść i odbiorca wizualizacji	12
1.4. Typ wizualizacji	18
2. Dobre praktyki tworzenia wizualizacji	25
2.1. Podstawowe zasady wizualizacji.....	25
2.2. Droga do wizualnej doskonałości	38
2.3. Katalogi wizualizacji i selektory wykresów	44
3. Manipulacje i przekłamanie na wykresach	65
3.1. Podatność wykresów na zniekształcenia.....	65
3.2. Zmiana skali	66
3.3. Perspektywa i kolorystyka.....	73
3.4. Obrót.....	75
3.5. Ukrywanie trendów	75
3.6. Dane procentowe zamiast danych bezwzględnych.....	77
3.7. Punkt odniesienia	78
3.8. Prezentowanie wybiórczych danych	79
3.9. Niedozwolone operacje na osi czasu histogramu.....	84
4. Badanie umiejętności studentów w zakresie wizualnego przekazu informacji.....	89
4.1. Cel badania	89
4.2. Metodyka badań	90

4.3. Analiza przygotowanych przez studentów wizualizacji	91
4.4. Informacje uzupełniające badanie	104
4.5. Podsumowanie wyników badań.....	107
Zakończenie.....	109
Bibliografia	111
Spis rysunków.....	118
Spis tabel.....	121
Spis wykresów	122
Streszczenie.....	124
Abstract.....	125

Wprowadzenie

Współcześnie, czyli w społeczeństwie informacyjnym i gospodarce opartej na wiedzy, szybki i nieograniczony dostęp do danych powoduje, że człowiek nie jest w stanie zapoznać się z nimi, a tym bardziej zweryfikować informacji, z nich wynikających. Rzeczywistość sprawia, że coraz bardziej cenione są wszelkie syntezy oraz inne zwarte i skondensowane formy prezentacji treści. Ludzie czerpią informacje, prześlizgując się wzrokiem po nagłówkach stron internetowych, zatrzymując się na moment na kolorowych grafikach i obrazach. Podobnie czynią menedżerowie. Zamiast opasłych analiz i sprawozdań pełnych liczb, żądają podsumowań i raportów z czytelnymi wykresami. Najważniejsze informacje przedstawione w formie graficznej często są podstawą podejmowania kluczowych decyzji biznesowych. Niniejsza monografia poświęcona jest dlatego zagadnieniom wizualizacji, które w dzisiejszym świecie stanowią bardzo ważną rolę w przekazie treści.

Pierwszy rozdział książki ma charakter teoretyczny. Dokonano w nim systematyzacji pojęć, wyjaśniono różnice między wizualizacją danych, informacji oraz wiedzy, a następnie przeanalizowano istotne elementy, które należy wziąć pod uwagę przy tworzeniu wizualizacji. Należą do nich: cel wizualizacji, jej treści, odbiorca oraz typ formy graficznej. Znajomość tych aspektów rzutuje na jakość tworzonych wizualizacji i tym samym wpływa na skuteczność przekazywania danych, informacji i wiedzy między ludźmi, co z biznesowego punktu widzenia często decyduje o trafności podejmowanych działań.

Pozostałe rozdziały mają charakter praktyczny. W drugim rozdziale zawarto wiele cennych wskazówek, dotyczących zasad tworzenia skutecznych wizualizacji, pochodzących od światowej klasy specjalistów w tym obszarze. Poruszono aspekty, wpływające na czytelność i jednoznaczność przekazu, a także na jego dostępność. Dokonano przeglądu wielu witryn i serwisów internetowych

poświęconych wizualizacjom, selektorom wykresów oraz narzędziom informatycznym wspierającym tworzenie grafik. W założeniu autorki, przytaczane w tym rozdziale linki do stron internetowych mają stanowić źródło inspiracji dla czytelników i przestrzeń do poszerzania wiedzy w zakresie wizualizacji.

W rozdziale trzecim przedstawiono możliwości wypaczania faktów i manipulacji przy przekazie informacji, jakie zdarza się stosować projektantom wykresów w celu wywołania u odbiorcy określonego sposobu myślenia o prezentowanej sytuacji. Przeanalizowano zarówno przekłamania wynikające ze świadomego wykorzystania funkcjonalności komputerowych programów graficznych stosowanych do tworzenia wykresów, jak i zniekształcenia będące efektem celowego doboru takich, a nie innych danych do wizualizacji.

W rozdziale czwartym z kolei przedstawiono wyniki badań przeprowadzonych wśród osób, którym z racji wykształcenia wizualizacja danych nie powinna być obca. Zbadano poziom ich umiejętności w zakresie tworzenia wykresów na podstawie opisu słownego. Eksperyment polegał na wykonaniu przez badanych wykresu na bazie krótkiego słownego opisu pewnej sytuacji. Analizie poddano zarówno formę wizualną, jak i zawartość merytoryczną wykresów. Pozwoliło to na identyfikację najczęstszych błędów popełnionych przez badanych, które zdecydowanie ograniczyły czytelność wizualizacji bądź wręcz sprawiły, że była ona niepoprawna lub jedynie fragmentaryczna.

Serdecznie zachęcam czytelników do zapoznania się z poszczególnymi rozdziałami, wynikami badań, a także z zawartością stron internetowych i portali poświęconych wizualizacji, do których linki podaję w tekście, gdzie znaleźć można wiele cennych rekomendacji dotyczących tworzenia efektywnych wizualizacji oraz informacji o możliwościach wykorzystania narzędzi informatycznych, które znacznie przyspieszą i ułatwią to zadanie.

Autorka

1. Elementy tworzenia użytecznych wizualizacji

1.1. Definicje wizualizacji

Wizualizacja jest zagadnieniem interdyscyplinarnym, będącym w kręgu zainteresowania specjalistów z różnych dziedzin i dyscyplin m.in. informatyki, a w szczególności grafiki komputerowej, zarządzania, dziennikarstwa, socjologii czy nauki o informacji. Naukowcy z różnych obszarów badają wizualizację z własnej perspektywy, zwracając uwagę na różne elementy (Motylińska, 2014). W niniejszej monografii wizualizacja jest analizowana głównie z punktu widzenia nauk o zarządzaniu.

Manovich (2010) zdefiniował wizualizację jako konwersję mierzalnych danych na ich wizualną reprezentację. Shirley i Marschner (2009) piszą, że wizualizacja to użycie obrazów do reprezentowania zmiennych przestrzennych i nieprzestrzennych w sposób, który ujawnia pewne prawidłowości i relacje. Dla Bertschi i in. (2011) wizualizacja to proces obejmujący różne etapy, takie jak gromadzenie, interpretacja, opracowywanie, rozumienie, projektowanie i udostępnianie informacji. Card i in. (1999) podają definicję wizualizacji informacji jako wykorzystania wspomaganych komputerowo, interaktywnych, wizualnych reprezentacji abstrakcyjnych danych w celu wzmocnienia funkcji poznawczych. Z kolei Yusoff i in. (2013) odnoszą się do wizualizacji wiedzy, pisząc, że polega ona na przedstawianiu złożonych koncepcji i danych za pomocą grafiki i animacji w sposób, jakiego ludzie wcześniej nie widzieli, w celu ułatwienia transferu i tworzenia nowej wiedzy.

Jak widać, wizualizacja może być reprezentacją danych, informacji lub wiedzy. Za każdym razem odnosi się do innego poziomu abstrakcji (Chen i in., 2009), zgodnie z definicjami pojęć (Krok, 2013):

- dane – to wszelkie nieuporządkowane, pojedyncze, niepołączone ze sobą fakty o jakimś wydarzeniu bądź obiekcie,
- informacja – to dane, którym nadano znaczenie,

- wiedza – powstaje po skojarzeniu dostępnych informacji, umieszczeniu ich w kontekście i wyciągnięciu wniosków w celu podjęcia decyzji.

Można zatem przyjąć, że wizualizacja danych pociąga za sobą każdy wizualny artefakt, który wyjaśnia dane w dowolnej dyscyplinie (Gatto, 2015) i ujawnia szczegóły, które mogły zostać ukryte w obliczonych statystykach (Nielsen, 2016). Jest graficzną ilustracją (statyczną lub interaktywną) pewnych istotnych cech zbioru danych (zmierzonych wielkości odnoszących się do jakiegoś zagadnienia). Dotyczy przeważnie dużych zbiorów danych, często tworzona jest automatycznie, a jej celem jest określanie ważnych cech i wzorców w zbiorze danych z zamysłem późniejszych analiz statystycznych i podejmowania decyzji. Pełni zatem rolę głównie eksploracyjną. Pozwala projektantowi (analitykowi) badać, modyfikować, rozważać i wypróbowywać różne scenariusze, typu „co by było gdyby”. Co ważne, wizualizacja danych ma zapewnić obiektywizm i umożliwić odbiorcom samodzielne wyciąganie wniosków.

Wizualizacja informacji wynika natomiast ze wstępnych analiz danych i wyraża zamiar wpłynięcia projektanta na odbiorcę, doprowadzenia go do konkretnych (a zatem subiektywnych) wniosków. Najczęściej tworzona jest dla określonego zbioru danych, w celu zwrócenia uwagi odbiorcy na te relacje, na których pokazaniu zależy projektantowi. Prezentuje zatem przemyślaną historię, która prowadzi odbiorców do „zaplanowanych” wniosków. Można ją więc traktować jako pewien przekaz, infografikę, formę komunikacji między projektantem wizualizacji a odbiorcą.

Różnice między wizualizacją informacji a wizualizacją wiedzy wyjaśniają m.in. Burkhard (2005) oraz Van Biljon i Renaud (2015), pisząc, że ta pierwsza ma głównie wspierać identyfikację wzorców i odkrywanie prawidłowości, a celem tej drugiej jest przede wszystkim transfer wiedzy. Można zatem powiedzieć, że wizualizacja informacji poprawia dostęp do informacji, pozwala uzyskać nowe spojrzenie na fakty, pomaga użytkownikom zrozumieć treść, a wizualizacja wiedzy usprawnia procesy komunikacyjne i ułatwia jednostkom lub grupom osób przekazywanie wiedzy i tworzenie nowej wiedzy w warunkach współpracy.

Zhang i in. (2008) pod pojęciem wizualizacji wiedzy rozumieją czynność polegającą na badaniu wykorzystania reprezentacji wizualnych, takich jak wykresy, diagramy, rysunki, w celu usprawnienia tworzenia wiedzy i transferu między co najmniej dwiema osobami. Eppler (2011) odnosi termin wizualizacji do wykorzystania grafiki w celu tworzenia, integrowania i administrowania wiedzą.

Meyer (2010) twierdzi, że wykorzystanie jedynie tekstowych reprezentacji wiedzy nie spełnia wymagań współczesnego społeczeństwa informacyjnego. Popierają

go Kelleher i Wagener (2011), którzy uważają, że wizualizacja ma zdolność syntezy danych w efektywną grafikę, ułatwiając w ten sposób ludzkiemu mózgowi zrozumienie danych. Szews (2017) dodaje, że przekaz wizualny jest łatwiej dostrzegalny i ma większą siłę perswazji niż zwykły tekst, ponieważ mocno oddziałuje na bodźce wzrokowe. Yaacob, Liang i Mohamad (2017) wskazują, że wizualizacja jest środkiem wspomagającym funkcjonowanie zmysłów u ludzi. Burkhard (2005) uzasadnia skuteczność wizualizacji, podając kilka potwierdzonych w literaturze faktów:

- pojemność kanału wejściowego człowieka jest większa, gdy wykorzystywane są zdolności wzrokowe (Miller, 1956),
- ludzkie mózgi mają silną zdolność rozpoznawania wzorców (Ellis, 1938; Koffka, 1935),
- przywołanie wizualne u ludzi wydaje się być skuteczniejsze niż werbalne, ludzie mają wrodzone zdolności do efektywnego przetwarzania reprezentacji wizualnych (Kosslyn, 1980; Shepard, Cooper, 1982),
- badania empiryczne wykazały, że w różnych zadaniach reprezentacje wizualne są lepsze od reprezentacji werbalno-sekwencyjnych (Larkin, Simon, 1987; Glenberg, Langston, 1992; Bauer, Johnson-Laird, 1993; Novick, 2001).

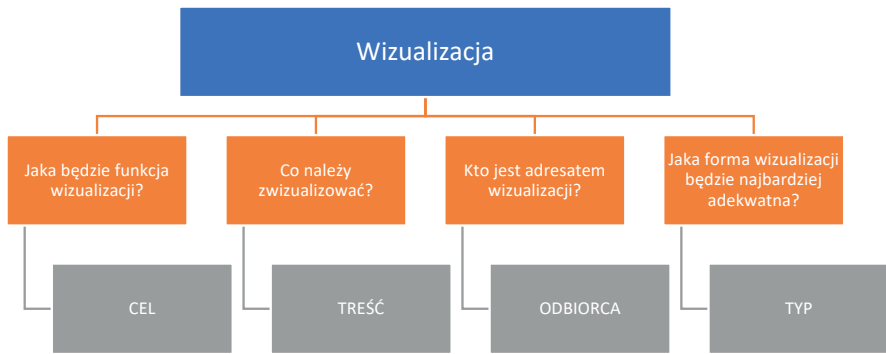
Zagadnieniom przekazu graficznego poświęcono wiele prac naukowych. Biećek (2014) w swoich esejach systematyzuje wiedzę przydatną do projektowania i wykonania dobrej, czytelnej i estetycznej wizualizacji. Midway (2020) udziela cennych wskazówek dotyczących sztuki prezentowania danych oraz podaje podstawowe zasady efektywnych i skutecznych wizualizacji danych. Podobnie Wilke (2020) w swojej książce przedstawia reguły pomocne w projektowaniu estetycznych wykresów oraz w doborze technik wizualizacji do różnych rodzajów danych. Gatto (2015) omawia potrzeby, wyzwania i przedstawia dobre praktyki w obszarze tworzenia wizualizacji, szczególnie z punktu widzenia poprawy skuteczności udostępniania wyników badań akademickich opinii publicznej, decydom politycznym, mediom, przedsiębiorstwom i innym zainteresowanym stronom. Bresciani i Eppler (2015) z kolei analizują typowe pułapki oraz potencjalne wady i ograniczenia reprezentacji treści w formie graficznej. Rośnie też grono specjalistów-praktyków prowadzących witryny internetowe i realizujących projekty związane z różnymi aspektami wizualizacji, o czym mowa będzie w kolejnym rozdziale książki.

Im więcej danych, informacji i wiedzy, tym trudniejsza eksploracja, odczyt, zrozumienie, przekazanie i wyciągnięcie wniosków. Wizualizacja ma moc ujawniania wzorców, prawidłowości, luk, wartości odstających w danych, potrafi znacznie zwiększyć efektywność transferu wiedzy, a obecne technologie informatyczne

pozwalają na dobór optymalnej, skutecznej, atrakcyjnej a nawet interaktywnej formy przekazu.

Powyższa analiza wskazuje, że ostateczny rezultat wizualizacji koncentruje się na użytkowniku, który postrzega treść, niezależnie od tego, czy są to dane, informacje czy wiedza.

Aby stworzyć użyteczną wizualizację, należy odpowiedzieć sobie na cztery istotne pytania, które pozwolą ustalić cel, treść, odbiorcę i najlepszy w danym przypadku typ wizualizacji (rys. 1.1).



Rysunek 1.1. Elementy, które należy przeanalizować przed rozpoczęciem wizualizacji

Źródło: opracowanie własne.

Poszczególne elementy wizualizacji szczegółowo omówiono w kolejnych podrozdziałach.

1.2. Cel wizualizacji

Jeśli chodzi o rolę i funkcję wizualizacji, to przegląd literatury przedmiotu (Burkhard, 2005; Roberts i in., 2014; Valkanova, Jorda, Vande Moere, 2015; Bresciani, Eppler, 2015; Renaud, Van Biljon, 2017) oraz własne przemyślenia autorki pozwoliły zidentyfikować następujące cele wizualizacji:

1. Pokazanie zależności – przede wszystkim przy wizualizacji danych lub informacji, ukazanie relacji między seriami lub punktami danych, wyeksponowanie niewidocznych na pierwszy rzut oka związków.
2. Odkrycie trendów, prawidłowości, wzorców i wyjątków – podobnie jak w przypadku ukazania zależności, odkrywanie trendów dotyczy głównie wizualizacji

danych lub informacji; dzięki graficznej prezentacji łatwiejsze staje się dostrzeżenie pewnych prawideł, tendencji, orientacji, odstających/niepasujących danych lub dostrzeżenie luk.

3. Tworzenie i wywoływanie nowych spostrzeżeń – wizualizacje poprzez konieczność samodzielnej interpretacji przez użytkownika tworzą warunki do kreowania wiedzy, do powstawania nowej wiedzy, do konwersji wiedzy dostępnej w wiedzę ukrytą, czyli do procesu internalizacji na „spirali wiedzy” Nonaki i Takeuchiego.
4. Skodyfikowanie, czyli odpowiednie zapisanie, zapamiętanie w adekwatnej formie, przelanie na papier lub na postać elektroniczną w celu późniejszego udostępniania/wykorzystywania, eksternalizacji wiedzy.
5. Pokazanie jednocześnie ogółu i szczegółów – dobrym przykładem są mapy wiedzy, które można potraktować jak graficzne katalogi źródeł wiedzy, zasobów, struktur lub procesów. Obejmują pewien obszar, wskazują kontekst, pokazują zarówno widok ogólny (przegląd), jak i dają wgląd w szczegóły oraz wzajemne powiązania między nimi.
6. Podzielenie się wiedzą/transfer wiedzy – wizualizacja może być po prostu sposobem na przekazanie myśli, odczuć, doświadczeń bądź spostrzeżeń, próbą socjalizacji, wspomagającą zamianę wiedzy ukrytej nadawcy w wiedzę ukrytą odbiorcy.
7. Ułatwianie współpracy – reprezentacje wizualne ułatwiają procesy komunikacyjne w zespołach, np. poprzez wyrażanie odczuć, popieranie pomysłów za pomocą emotikonów, podnoszenie ręki w głosowaniach lub w celu zabrania głosu.
8. Wspieranie zapamiętywania i przypominania sobie – wszelkie wizualizacje towarzyszące słownemu opisowi rzeczywistości lub przekazowi wiedzy przyczyniają się do lepszego przyswojenia, zrozumienia i tym samym zapamiętania treści.
9. Ułatwienie nauki – poprzez angażowanie zmysłu wzroku wizualizacje wspierają rozumowanie i komunikację, oddziałują na wyobraźnię, mają potencjał wyrażania trudnych i skomplikowanych treści, często wywołują efekt „aha”.
10. Wywołanie emocji – prezentacje wizualne przedstawiające rzeczywistość m.in. w sposób ekspresyjny, inspirujący, motywujący, których celem jest pobudzenie odbiorcy, zmuszenie do refleksji, wywołanie określonego uczucia, np. przeżycia, ekscytacji, strachu, podziwu; funkcja często spotykana w reklamie i marketingu.

11. Zdobycie i utrzymanie uwagi odbiorców – przykucie uwagi i dłuższe skupienie słuchacza/widza na przekazywanych treściach może być bardzo istotne dla nadawcy informacji, zwłaszcza przy omawianiu/prezentacji ważnych i jednocześnie trudnych zagadnień.
12. Motywowanie, inspirowanie, aktywizowanie odbiorców – odpowiednio zaprojektowane reprezentacje wizualne skłaniają do myślenia i tym samym zaangażowania się w temat, zachęcają odbiorcę do pogłębiania wiedzy i motywują do działania.

Określenie celu wizualizacji w znacznym stopniu rzutuje na dobór treści przekazu oraz adekwatnej formy graficznej.

1.3. Treść i odbiorca wizualizacji

Kolejne pytanie, na jakie trzeba odpowiedzieć to: „co będzie wizualizowane?” Dane, informacje czy może wiedza? Jeśli dane, to które, jeśli wiedza, to jaka?

W przypadku danych należy wziąć pod uwagę skalę pomiaru i typ zmiennych. Elementy te decydują przede wszystkim o możliwych do zastosowania metodach statystycznych podczas prowadzonych analiz, ale świadomość rodzaju zmiennych jest istotna również z punktu widzenia wizualizacji i doboru odpowiedniej metody, np. właściwego typu wykresu.

W teorii pomiaru wyróżnia się cztery podstawowe skale, wprowadzone przez Stevensa (1959):

- nominalną – na której między wartościami zmiennych zachodzą tylko relacje równości lub różności.
- porządkową – liczby przynoszą informacje jedynie o kolejności w szeregu danych obiektów ze względu na mierzoną cechę, na wartościach tych można wyznaczać relacje równości, różności, większości i mniejszości,
- interwałową (przedziałową) – oprócz kolejności w szeregu, liczby zawierają również informacje o wielkości różnic między obiektami w zakresie mierzonej cechy, stąd możliwe są relacje równości, różności, większości i mniejszości, a także równości różnic i przedziałów,
- ilorazową (stosunkową) – liczby dostarczają informacji nie tylko o uporządkowaniu obiektów i względnej wielkości różnic, lecz również o relacji między wielkościami, stąd oprócz relacji równości, różności, większości, mniejszości, równości różnic i przedziałów, możliwe są relacje równości stosunków między poszczególnymi wartościami skali.

Dwie pierwsze to skale słabe, dwie następne – skale mocne. Cechy mierzone na skali nominalnej nazywa się cechami jakościowymi lub kwalitatywnymi, a mierzone na pozostałych skalach – ilościowymi lub kwantytatywnymi, choć niektórzy autorzy, jak na przykład Steczkowski i Zeliaś (1981, s. 22), do cech jakościowych zaliczają również te mierzone na skali porządkowej.

Zmienne na skali nominalnej grupują obiekty według wartości, jaką przyjmuje zmienna dla obiektu. Możliwe jest porównywanie wartości zmiennej między obiektami – równość lub różność pod względem zmiennej. Zmienne porządkowe pozwalają już na uporządkowanie obiektów – można stwierdzić, któremu obiektowi przysługuje zmienna porządkowa w większym stopniu. Jeśli na przykład do opisu natężenia jakiejś cechy użyto skali porządkowej i wyrażono je za pomocą liczb całkowitych z zakresu od 1 do 6, to liczby te przynoszą jedynie informacje o kolejności w szeregu danych obiektów ze względu na mierzoną cechę. Jeśli natężenie danej cechy w obiekcie X wynosi 6, a w obiekcie Y wynosi 2, oznacza to jedynie tyle, że natężenie tej cechy w obiekcie X jest większe niż jej natężenie w obiekcie Y. Nie ma jednak żadnej informacji o względnej wielkości różnic ani o relacjach między wielkościami. Nie można zatem powiedzieć, że natężenie tej cechy w pierwszym obiekcie jest o 4 większe niż w drugim, ani 3 razy większe niż w drugim. Różnice między dwoma zmiennymi można obliczać dopiero, gdy pomiaru dokonano na skali interwałowej (pomiar temperatury, dat, szerokości/długości geograficznej). Zmienne ilorazowe z kolei pozwalają stwierdzić, że natężenie danej zmiennej dla obiektu X jest k razy większe/mniejsze niż tej samej zmiennej dla obiektu Y.

Typ skali, ze względu na dopuszczalne przekształcenia, determinuje przede wszystkim stosowalność technik statystyczno-ekonometrycznych. W przypadku mierzenia zmiennych na skalach wyższego rzędu, czyli przedziałowych lub ilorazowych wybór technik statystycznych jest znacznie szerszy niż dla pomiaru dokonywanego na skalach słabych, tj. nominalnej lub porządkowej.

Warto też mieć świadomość typu analizowanych i prezentowanych zmiennych. W statystyce rozpatruje się dwa rodzaje zmiennych – ciągłe i dyskretne. Zmienna ciągła to taka, która przyjmuje nieprzeliczalną liczbę wartości z określonego przedziału, natomiast zmienna dyskretna – skończoną lub przeliczalną.

Z punktu widzenia prowadzenia badań eksperymentalnych, a także wizualizacji ich wyników ważny jest również podział na zmienne niezależne (inaczej objaśniające) i zależne (inaczej objaśniane). Zmienne niezależne to te, którymi manipuluje się w badaniu (eksperymentcie), tj. zmienia się np. ich wartości, kierunek lub siłę. Z kolei zmienne zależne są tymi, które podlegają pomiarom, a ich wartość jest „zależna” od wartości zmiennej niezależnej. W badaniu mierzy się wpływ

zmiennych niezależnych na zmienne zależne. Dla przykładu, gdy badacz chce ustalić, czy mężczyźni różnią się od kobiet pod względem sprawności fizycznej, to zmienną niezależną w opisywanym przypadku jest płeć, a sprawność fizyczna zmienną zależną. Gdy chce zbadać, czy wykształcenie ludzi wpływa na liczbę języków obcych, którymi umieją się posługiwać, zmienną niezależną będzie wykształcenie, a znajomość języków obcych zmienną zależną.

Wiedza i rozumienie typu danych oraz skal, na jakich dokonano pomiarów może pomóc w wyborze optymalnej formy graficznej, jako że różne metody wizualizacji w różnym stopniu nadają się do pokazania danych w zależności od ich rodzaju. Czapiewski (2010) na swoim blogu omówił możliwości wyboru prawidłowych form wizualizacji w zależności od prezentowanych związków między danymi, natomiast Stanimir (2012) przedstawiła różne techniki prezentacji powiązań kategorii zmiennych nominalnych.

Na dobór odpowiedniej formy graficznej, zwłaszcza przy prezentacji informacji, ma również wpływ typ zagadnienia, które projektant chce przekazać. Chodzi o zadanie, jakie twórca wizualizacji stawia przed odbiorcą: czy ma to być porównanie wartości ze sobą, czy zestawienie służące pokazaniu, jaki jest udział określonej części w całości, czy może celem jest zwrócenie uwagi na rozkład i częstotliwość występowania pewnych wartości w zestawie danych, a może też intencją projektanta jest skupienie uwagi odbiorcy na pewnych relacjach między prezentowanymi danymi. Większość narzędzi, tzw. selektorów wykresów, które omawiane będą w kolejnym rozdziale książki, pozwala na filtrowanie zbioru wykresów pod takim właśnie kątem – ich przydatności do prezentacji danego typu zagadnienia. Niektóre z selektorów uwzględniają również liczbę i rodzaj zmiennych jako kryterium doboru optymalnej formy graficznej.

Jeśli chodzi o wizualizację wiedzy, to w literaturze przedmiotu autorzy (m.in. Burkhard, 2004; Renaud, van Biljon, 2017) najczęściej przedstawiają klasyfikację wiedzy odpowiadającą na pytania: co (wiedza deklaratywna), jak (proceduralna), dlaczego (eksperymentalna), gdzie (orientacyjna) i kto (indywidualna).

Autorka proponuje poszerzenie zestawu pytań, na jakie może odpowiadać wizualizowana wiedza i ich podział, zgodnie z typem wiedzy, na:

- wiedzę jawną (*explicit*) odpowiadającą na pytania: co, kto, kiedy, gdzie, czyli mającą charakter bardziej informacyjny, łatwą do kodyfikacji, gromadzenia, przechowywania i tym samym wizualizacji,
- wiedzę ukrytą (*tacit*), odpowiadającą na pytania: dlaczego, co jeśli, czy gdyby, w jaki sposób, czy można, po co, skąd się bierze, co zrobić, żeby, trudną do kodyfikacji i przekazania.

Burkhard (2004) zwraca uwagę, że nie istnieją jednoznaczne wytyczne, które wskazywałyby, jaki typ wizualizacji jest odpowiedni do przekazania danego typu wiedzy. W praktyce bardzo często wizualizacje tworzone są na potrzeby konkretnych odbiorców i wtedy przekazywana treść powinna zależeć od potrzeb użytkownika grafiki. Odpowiedź na pytanie, co należy zwizualizować, jest zatem ściśle powiązana z odpowiedzią na kolejne pytanie – kto jest odbiorcą wizualizacji?

Burkhard (2004) i Renaud oraz van Biljon (2017) wyróżniają odbiorców indywidualnych, grupy, zespoły, konkretne organizacje. Podkreślają wagę różnic w percepcji i zdolności do zrozumienia, wynikające m.in. z wieku, wykształcenia i dotychczasowego doświadczenia odbiorców. Wizualizacja powinna być bowiem dostosowana do umiejętności postrzegania użytkowników. Znajomość kontekstu i poznawczego tła grupy docelowej będą bardzo pomocne przy wyborze właściwej metody wizualizacji.

Należy jednak zwrócić również uwagę na potrzeby i oczekiwania odbiorców w zakresie wizualizowanych treści. Przez kogo wizualizacja będzie wykorzystywana, jak będzie doświadczana, czy spełni potrzeby użytkownika od strony funkcjonalności i użyteczności? Temu zagadnieniu poświęcono nieco więcej miejsca poniżej, omawiając je na przykładzie. Załóżmy, że mamy zbiór danych dotyczących wyposażenia, a dokładnie mebli będących na stanie pewnego klubu seniora. Zbiór dostępny jest w układzie tabelarycznym (zatem jest już w pewien sposób zwizualizowany) i zawiera takie kolumny jak: identyfikator produktu, nazwa produktu, producent, liczba sztuk na stanie, data zakupu, cena oraz wartość (tab. 1.1).

Tabela 1.1. Dane prezentujące wyposażenie klubu seniora

Identyfikator produktu	Nazwa produktu	Producent	Liczba sztuk na stanie	Data zakupu	Cena (zł)	Wartość (zł)
KRZE430	krzesło	Melbest	30	20.07.2016	120	3 600
STO231	stół	Faber	17	12.09.2016	250	4 250
SZAF983	szafka	Verdi	15	14.10.2016	200	3 000
REG67	regał	Profi	15	05.10.2018	100	1 500
TAB893	taboret	Profi	10	20.09.2018	70	700
FOT983	fotel	Verdi	15	23.04.2018	300	4 500
SZAF392	szafka	Faber	2	15.02.2019	180	360
SZAF129	szafka	Faber	8	15.02.2019	170	1 360
REG983	regał	Faber	10	17.05.2019	110	1 100
KRZE781	krzesło	Profi	35	20.01.2021	150	5 250

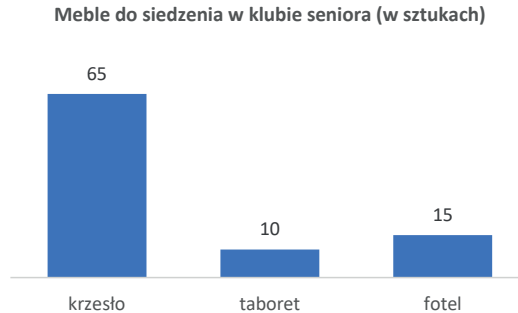
Źródło: opracowanie własne.

Tabela zawiera zbyt dużo danych, by sensownie można było zaprezentować je na jednej grafice. Wykres powinien przyjąć formę i zakres dostosowane do odbiorcy wizualizacji i jego potrzeb. Gdy odbiorca oczekuje konkretnej informacji, to taką należy mu dostarczyć. Nie mniej, ale też nie więcej niż potrzebuje w danym momencie. Nieistotna zawartość komunikatu rozprasza i wydłuża czas odbioru.

Wracając do przykładu z meblami, inny będzie zakres i postać wizualizacji, gdy jej odbiorcą ma być organizator spotkania dla seniorów, którego interesuje liczba miejsc siedzących, jaką może zaoferować uczestnikom spotkania. W innym ujęciu i inne dane należy pokazać, gdy odbiorcą jest osoba z działu finansowego klubu, którą interesuje wartość poszczególnych produktów łącznie z danymi o ich identyfikatorze i jeszcze w innym ujęciu, gdy odbiorcą informacji jest osoba odpowiedzialna za wyposażenie klubu, która chce mieć zwięzły raport, z którego łatwo odczyta od jakich producentów i jakie meble kupuje.

By przekazać interesujące odbiorców informacje, dane należy przefiltrować i przedstawić w odpowiednim ujęciu, w zależności od potrzeb użytkownika.

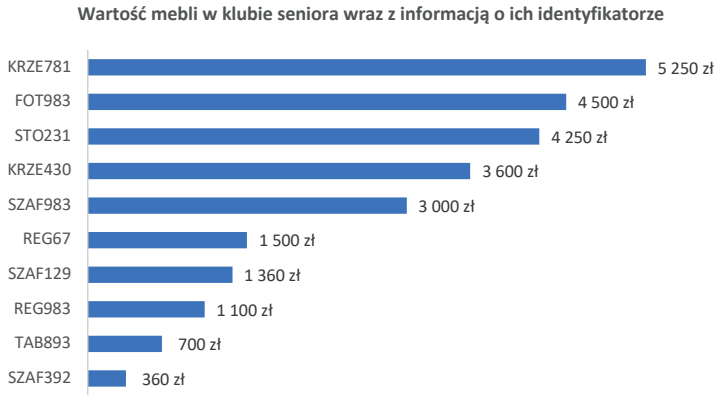
Dla pierwszego odbiorcy – organizatora spotkania dla seniorów – odpowiednia będzie następująca wizualizacja (wykres 1.1):



Wykres 1.1. Wizualizacja zgodna z potrzebami organizatora spotkań dla seniorów

Źródło: opracowanie własne.

Drugi odbiorca – pracownik działu finansowego klubu – interesujące go informacje łatwo odczyta z następującej wizualizacji (wykres 1.2):

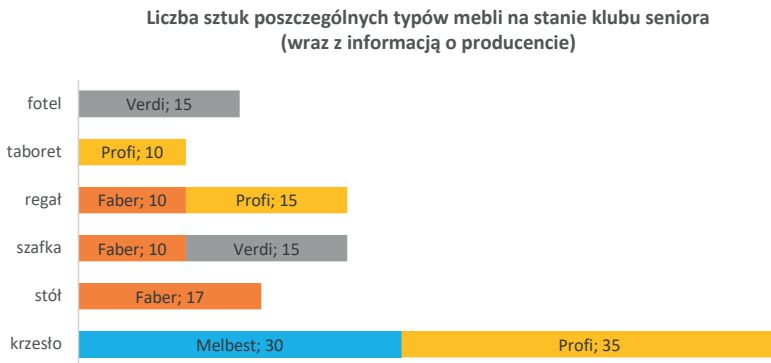


Wykres 1.2. Wizualizacja zgodna z potrzebami pracownika działu finansowego klubu seniora – sortowanie według wartości

Źródło: opracowanie własne.

W tym przypadku rozważyć należy również odpowiednie posortowanie danych, w zależności od tego czy odbiorca potrzebuje zestawienia mebli według ich wartości czy może raczej według identyfikatora mebla. Informacje powinny być mu zaprezentowane w układzie jak najbardziej wygodnym do dalszych analiz i podejmowania decyzji.

Trzeciemu odbiorcy zaś – osobie odpowiedzialnej za wyposażenie klubu seniora – pomocna będzie następująca wizualizacja (wykres 1.3):



Wykres 1.3. Wizualizacja zgodna z potrzebami pracownika odpowiedzialnego za wyposażenie klubu seniora

Źródło: opracowanie własne.

Przy czym, w tym przypadku zdecydowanie najlepiej byłoby mieć do dyspozycji wizualizację interaktywną, na przykład w postaci „kokpitu menedżerskiego”, co pozwoliłoby na filtrowanie danych według typu mebla oraz producenta i wyświetlanie tylko tych, które w danym momencie pracownik chce przeanalizować.

Dla żadnego z odbiorców wizualizacji nie jest istotna data zakupu ani cena za sztukę danego mebla, stąd nie ma sensu prezentować ich na wykresie.

1.4. Typ wizualizacji

Czwartym aspektem, któremu należy poświęcić uwagę przy projektowaniu wizualizacji jest jej typ. Wyróżnić można wiele form i metod wizualizacji. Każdy typ wizualizacji ma określone cechy oraz mocne i słabe strony. Poniżej dokonano zestawienia i krótkiej charakterystyki wybranych typów.

Szkice – przedstawiają główną ideę, są szybkie w realizacji, używa się ich do wspomagania grupowej refleksji i procesu komunikacji, na przykład podczas warsztatów, wspierają prowadzenie dyskusji, argumentowanie i wymianę wiedzy. Jeśli powstają w procesie przekazywania wiedzy, to dodatkowo podtrzymują uwagę i skupienie odbiorcy, pozostawiają miejsce na własne interpretacje i sprzyjają kreatywności. Można je łatwo przerabiać, wzbogacać, doprecyzowywać, dostosowując do potrzeb (zwłaszcza, gdy powstają w bieżącym kontakcie z odbiorcą).

Diagramy/schematy – służą uproszczonemu przedstawieniu abstrakcyjnych idei, najczęściej wykorzystują znormalizowane kształty do uporządkowania informacji, zilustrowania funkcji poszczególnych elementów i relacji między nimi. Mają charakter analityczny, systematyzujący lub klasyfikujący. Ich głównym celem jest prezentacja pewnego modelu i rodzaju zależności między jego elementami. Gdy celem wizualizacji jest transfer wiedzy, to diagramy na pewno czynią abstrakcyjne pojęcia bardziej dostępnymi dla odbiorcy, pomagają zmniejszyć ich złożoność, wzmocnić poznanie, uporządkować strukturę informacji i wyjaśnić związki przyczynowe. Podstawowy zestaw schematów oferuje na przykład SmartArt w ramach pakietu MS Office (rys. 1.2).



Rysunek 1.2. Zestaw schematów w pakiecie MS Office: lista, proces, cykl, hierarchia, relacja, macierz, piramida

Źródło: wycinki ekranu z ikonami grafik SmartArt pakietu MS Office.

Wykresy – to analityczne narzędzie graficzne, odpowiednie do prezentacji zmienności zjawiska w czasie, porównania kilku wielkości, pokazania zależności lub zobrazowania proporcji. Typ wykresu (rys. 1.3) dobierany jest pod kątem rodzaju danych lub informacji, które będą wizualizowane, na przykład wykresy słupkowe/kolumnowe w czytelny sposób ukazują różnice między zmiennymi jakościowymi. Paski wykresu słupkowego/kolumnowego można dowolnie przedstawiać (np. alfabetycznie, według długości słupka), co pomaga użytkownikowi w lepszej i szybszej interpretacji przedstawionych danych. Innym popularnym rodzajem wykresów są histogramy, które służą do prezentacji rozkładów zmiennych. Przedstawiają dane ilościowe z zakresami danych pogrupowanymi w przedziały. Wyglądem przypominają wykresy słupkowe, jednak układu pasków w histogramie nie można zmieniać, bo traci on sens. Wartości na histogramach reprezentowane są przez przylegające do siebie obszary. Popularnym, zwłaszcza wśród analityków oraz inwestorów, typem wykresu są też wykresy liniowe, stosowane na przykład do ukazywania zmienności kursów walutowych lub do szybkiego rozpoznania tendencji danej zmiennej w badanym okresie. Podstawą ich popularności jest przejrzystość oraz łatwość interpretacji. Do wykresów liniowych podobne są wykresy warstwowe. Warstwy poniżej linii wypełnia się kolorami lub teksturami. Używane są przede wszystkim do przedstawiania wielkości zmiany w czasie. Główny nacisk kładzie się na podkreślenie różnic między kategoriami. Do obrazowania relacji części do całości wykorzystuje się skumulowane wykresy warstwowe. Często spotykanym (i równie często krytykowanym, o czym będzie mowa w kolejnym rozdziale) jest wykres kołowy, wykorzystywany w celu prezentacji udziału danego elementu bądź elementów jednej zmiennej w ich sumie, którą traktuje się jako 100%. Stosowany jest do przedstawienia proporcji, czyli udziału danej cechy (zbiorowości, kategorii) w całości. Jeszcze innym typem jest wykres punktowy nazywany też wykresem rozrzutu, rozproszenia, diagramem korelacji, czy wykresem XY. Pokazuje zależność między dwoma zmiennymi, jej kierunek i siłę. Dane są prezentowane jako zbiór punktów, w których wartość jednej zmiennej określa położenie

punktu na osi poziomej, a wartość drugiej zmiennej określa położenie punktu na osi pionowej. Skala przedstawianych danych w tej wizualizacji nie musi rozpoczynać się od punktu zerowego, ponieważ odzwierciedla zawsze wartości przypisane do danej zmiennej. Wizualizacja ta pomaga w odkrywaniu relacji między opisywanymi wartościami lub wykazywaniu braku tych relacji i pokazuje możliwe punkty odstające oraz uwidacznia trendy. Korelacje w tym typie wykresów nie zawsze muszą być liniowe, a punkty danych nie muszą być rozłożone równomiernie. Wykresy punktowe często stosowane są przez naukowców, analityków, inżynierów oraz statystyków. Charakterystyczne dla tego typu wykresów jest to, że im więcej danych, tym lepiej widać zależności. Warto też omówić wykres radarowy, który umożliwia porównanie zagregowanych wartości kilku serii danych (obserwacji wielowymiarowych). Każda seria ma na wykresie unikatowy kolor lub wzór i jest reprezentowana w legendzie wykresu. Na jednym wykresie można wykreślić jedną lub kilka serii danych.



Rysunek 1.3. Przykładowe rodzaje wykresów w pakiecie MS Office: kolumnowy, słupkowy, liniowy, warstwowy, histogram, kołowy, punktowy

Źródło: wycinki ekranu z ikonami wykresów pakietu MS Office.

Programy informatyczne (typu arkusze kalkulacyjne, narzędzia statystyczne do analizy danych) oferują znacznie więcej typów wykresów, które nie zostały tu szczegółowo omówione, np. wykres bąbelkowy, giełdowy, powierzchniowy, pierścieniowy, kaskadowy, lejkowy i wiele innych (<https://support.microsoft.com/pl-pl/office/dost%C4%99pne-typy-wykres%C3%B3w-w-pakiecie-office-a6187218-807e-4103-9e0a-27cdb19afb90>). Obszerne informacje na temat możliwości zastosowania wykresów sekwencyjnych, skategoryzowanych, obrazkowych, macierzowych wraz z przykładami dostępne są na stronie firmy StatSoft (<https://www.statsoft.pl/textbook/>). Temat narzędzi informatycznych do tworzenia wykresów będzie też poruszony w kolejnym rozdziale książki, przy okazji omawiania portali poświęconych tematyce wizualizacji.

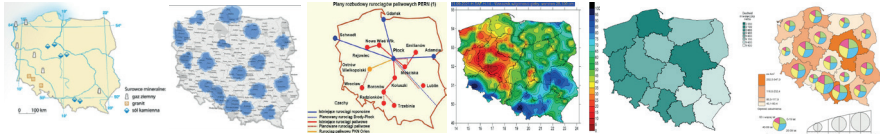
Tabele – równie często spotykane formy wizualizacji, składające się z wierszy i kolumn. Tabele pozwalają na prezentację wielu zbiorów wartości, porównanie ich poprzez sortowanie lub wybranie danych spełniających określone warunki poprzez

przefiltrowanie zgodnie z zadaniem kryterium. W tabelach można ukazać częściowe podsumowania i tym samym analizy na różnych poziomach szczegółowości. Dowolne przedstawianie wierszy i kolumn pozwala na taki dobór układu, aby w jak najbardziej czytelny sposób pokazać treści, które nadawca chce przekazać lub na które czeka odbiorca. Pola tabeli mogą zawierać zarówno teksty, liczby, daty, ale też symbole i obrazy.

Mapy tematyczne – to jedna z najpopularniejszych metod wizualizacji danych przestrzennych, mapy w czytelny i przyswajalny dla odbiorcy sposób prezentują „płaskie” dane powiązane z jednostkami geograficznymi. Ogólnie mówiąc, mapa jest graficznym, określonym matematycznie modelem rzeczywistości odniesionym do płaszczyzny, zgodnie z przyjętą skalą, prezentującą za pomocą symboli wartości atrybutów i związki przestrzenne między nimi (Medyńska-Gulij, 2011, s. 24). Na mapach można prezentować dane jakościowe odnoszące się do cech niemierzalnych, informujące jedynie o występowaniu danego zjawiska lub dane ilościowe, ukazujące natężenie danego zjawiska. Do najbardziej popularnych typów map tematycznych należą (rys. 1.4):

1. Mapa sygnaturowa (*signature map*) – polega na zastosowaniu na mapie symboli, tzw. sygnatur, w postaci odpowiednio dobranych znaków lub figur geometrycznych do przekazania informacji ilościowych lub jakościowych. Sygnatury mogą różnić się od siebie formą, barwą lub wielkością, mogą mieć formę figury geometrycznej, obrazka, litery lub symbolu kojarzącego się z przedstawianym obiektem.
2. Mapa zasięgu (*catchment map*) – również służy do prezentacji jednej zmiennej, np. obszaru jaki podlega pod dany urząd, rejonu o danym kodzie pocztowym, rejonu jaki obsługuje dana szkoła, strefy objętej płatnym parkowaniem, czyli pokazuje zasięg, jaki jest objęty daną cechą.
3. Mapa przepływu (*flow map*) – służy do prezentacji jednej zmiennej, pokazuje intensywność danego typu przepływu (trasy) przez odpowiednio dobraną szerokość linii. Przepływem mogą być na przykład najpopularniejsze kierunki podróży ludzi, przeloty samolotów (najczęstsze połączenia lotnicze), trasy żeglugi statków, transport towarów.
4. Mapa cieplna (*heat map*), inaczej mapa ciepła lub termiczna składa się z tzw. punktów ciepła – miejsc z największym nasileniem mierzonego zjawiska. Może być wykorzystywana do obserwowania zagęszczenia nasycenia obszaru daną usługą, produktem, do obserwowania nasycenia rynków. Podobną rolę pełni mapa choroplekowa (kartogram) z tym, że dotyczy granic geograficznych, tj. jednostek administracyjnych, takich jak państwo, województwo, gmina.

5. Kartodiagram – również popularny typ mapy tematycznej, przedstawiający zmienność wybranych cech obiektów przestrzennych za pomocą umieszczonych na mapie różnego typu wykresów. Ich lokalizacja odpowiada punktom pomiaru lub jednostkom przestrzennym, do których dane się odnoszą (np. gminom, województwom, państwom).



Rysunek 1.4. Przykładowe rodzaje map tematycznych: sygnaturowa, zasięgu, przepływu, termiczna, kartogram, kartodiagram

Źródło: wycinki ekranu z ikonami map odpowiednio ze stron: https://flipbook.nowaera.pl/dokumenty/Flipbook/Planeta-Nowa-7-podrecznik_2/index.html, <https://mvno-gsm.pl/dziesiec-nowych-nadajnikow-i-szesc-kolejnych-miast-objetych-technologie-dab/>, <https://www.kozieglowy.pl/aktualnosci/2367>, <http://nowoczesnafarma.pl/Problem-niedoboru-wody-nie-zniknal.html>, <https://predictivesolutions.pl/kartogram-w-poszukiwaniu-wzorcow-przestrzennych>, <https://pl.wikipedia.org/wiki/Kartodiagram>.

Obecnie można korzystać z wielu programów komputerowych wspierających zastosowanie tej formy wizualizacji. Udostępniają zaawansowane, intuicyjne, kontekstowe narzędzia kartograficzne, umożliwiając wizualizację danych na interaktywnych mapach. Zapewniają pełen repertuar symboli i stylów prezentacji kartograficznych. Oferują wiele map bazowych, które mogą być wykorzystane jako podstawa do prezentacji opracowywanej treści tematycznej. Odpowiednio dobrana mapa bazowa zapewnia kontekst, który ma kluczowe znaczenie dla zrozumienia prezentowanego zagadnienia. Można do nich dodawać wyskakujące okienka, wykresy i infografiki, wzbogacając w ten sposób przekaz informacji. Technologia IT daje ponadto możliwości wizualizacji danych w 3D, co pozwala jeszcze bardziej uatrakcyjnić przekaz informacji. Połączenie danych 2D i 3D nadaje im głębszy wymiar i pozwala pokazać pełny obraz z różnych perspektyw. Na przykład Produkty ArcGIS zapewniają wiele intuicyjnych narzędzi do tworzenia map i wizualizacji danych oraz wykonywania wszechstronnych analiz (<https://www.esri.pl/arcgis/>). Google również oferuje zaawansowane aplikacje do generowania map w różnych postaciach (*google earth, geochart, maps*). Możliwości trójwymiarowej wizualizacji danych daje również dodatek Microsoft 3D Maps w programie Excel. Jak podkreśla Król (2019) rozwój graficznych interfejsów obsługi oraz dostępność oprogramowania w oknie przeglądarki internetowej sprawiły, że interaktywne

wizualizacje danych stały się dostępne dla użytkowników bez specjalistycznej wiedzy. Do ich dyspozycji oddawanych jest coraz więcej narzędzi umożliwiających samodzielne tworzenie i publikowanie atrakcyjnych wizualizacji.

Obrazy – w przeróżnej postaci (zdjęć, ikon, rysunków, ilustracji, logotypów i innych) są bardzo atrakcyjną formą wizualizacji, zarówno w celu reprezentacji rzeczywistości, jak i transferu wiedzy, odnoszą się do emocji i inspirują; powodują wytwarzanie skojarzeń i wizji wartościowanych i przeżywanych przez człowieka, czasami wystarczy rzut oka na sugestywny obraz, by przekazywaną treść utrwalić na zawsze. Mózg ludzki jest w stanie szybko zapamiętać i równie szybko przywołać dany obraz, co może być bardzo pomocne do transferu, np. wiedzy biznesowej. Jak pisał Nonaka (1991) przekształcanie wiedzy ukrytej w wiedzę jawną oznacza znalezienie sposobu na wyrażenie tego, co niewyraźne i takim sposobem na wyartykułowanie doświadczeń lub spostrzeżeń może być właśnie użycie obrazu. Należy podkreślić, że obraz wraz z rozwojem nowych mediów (zdominowanych przez wizualizację) mocno ewoluował od formy analogowej i statycznej w stronę obrazu ruchomego, przestrzennego, cyfrowego i coraz częściej interaktywnego.

W zasadzie każdy z wymienionych typów wizualizacji może mieć **wirtualny charakter przestrzenny**, pozwalający na jeszcze pełniejsze i głębsze doświadczanie przez odbiorcę. Wymiar 3D nadawany zwłaszcza obrazom bądź mapom za pomocą technologii IT umożliwia ich obracanie i ogląd z różnych stron, zwiększa percepcję, ułatwia wyobrażenie sobie i tym samym zrozumienie idei i sensu przekazywanych treści.

Inna istotna cecha, która również może dotyczyć wszystkich analizowanych typów wizualizacji to ich **interaktywność**, zapewniana także dzięki odpowiedniemu oprogramowaniu. Interaktywność pozwala na zmianę wizualizowanych treści w zależności od potrzeb i jeszcze lepszą eksplorację danych i informacji. Chodzi tu o wszelkie możliwości filtracji tabel według zadanych kryteriów, zmianę dowolnych wielkości za pomocą suwaków i obserwacje zmieniających się danych, np. na wykresie lub mapie, jak również symulacje różnych wariantów i procesów w czasie oraz ich obserwacje na dynamicznie zmieniającej się wizualizacji. Wszystko to pozwala badać złożone dane, dokonywać nowych spostrzeżeń i tworzyć nową wiedzę.

W kolejnym rozdziale publikacji zawarto zbiór porad i wskazówek dotyczących tworzenia wizualizacji, a w szczególności prezentowania danych za pomocą wykresów. Zamieszczono w nim odsyłacze do wielu witryn prowadzonych przez specjalistów od wizualizacji oraz do platform oferujących katalogi z wykresami i możliwością ich filtracji, tak aby ułatwić dobór typu wykresu do danych. Omawiane witryny

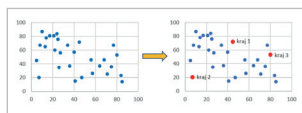
często zawierają wykaz i analizę różnorodnych narzędzi do tworzenia wykresów. W niektórych przypadkach bardziej szczegółowo zaprezentowano dane zagadnienie, w innych podano jedynie link do strony www, która może stanowić kolejne źródło poszerzania wiedzy i być inspiracją do dalszego zgłębiania tematu wizualizacji.

2. Dobre praktyki tworzenia wizualizacji

2.1. Podstawowe zasady wizualizacji

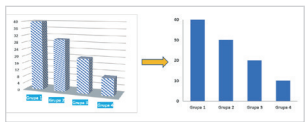
Jon Schwabish, autor kilku książek z zakresu percepcji wizualnej i przejrzystej komunikacji wyników badań, w 2018 roku na swoim blogu (<https://policyz.com/>) opublikował „miniściąawkę” dotyczącą podstawowych zasad wizualizacji danych (<https://policyz.com/2018/08/07/dataviz-cheatsheet/>). Post ten wyróżniono nagrodą KDnuggets Gold Blog (wiodącej witryny poświęconej uczeniu maszynowemu, sztucznej inteligencji i analityce) jako jeden z najlepszych blogów tamtego roku (<https://www.kdnuggets.com/2018/08/data-visualization-cheatsheet.html>). Poniżej przedstawiono kluczowe wskazówki Schwabisha, pochodzące z owego wpisu, a grafiki opracowała autorka niniejszej publikacji na bazie oryginałów pochodzących ze strony <https://policyz.com/2018/08/07/dataviz-cheatsheet/>.

POKAŻ ISTOTNE DANE



Wykres w raporcie badawczym, artykule lub blogu ma pomóc zrozumieć przekaz. Dane są najważniejszą częścią wykresu i powinny być przedstawione w możliwie najbardziej przejrzysty sposób. Co nie znaczy, że wszystkie dane muszą być pokazane – w praktyce wiele wykresów prezentuje zbyt wiele danych.

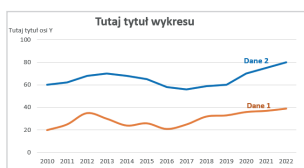
OGRANICZ LICZBĘ ROZPRASZACZY



„Bałagan” na wykresie, spowodowany nadmiernym i w zasadzie niczemu niesłużącym upiększeniem elementów wykresu poprzez ich przesadne sformatowanie sprawia, że przekaz wizualny jest mniej skuteczny.

Niepotrzebny trzeci wymiar, zaakcentowane podłoże, gęste i wyraźne linie siatki, gradientowe wypełnienie kolumn i ozdobne podpisanie kategorii znacznie osłabiają czytelność grafiki.

ZINTEGRUJ TEKSTY Z WYKRESEM



Standardowe raporty badawcze często konstruowane są jak pokaz slajdów, w którym autor opowiada to, co pojawia się na slajdzie. Lepszy model to taki, w którym wizualizacje konstruowane są tak, aby uzupełniały tekst, a jednocześnie zawierały wystarczającą liczbę informacji, by istnieć samodzielnie.

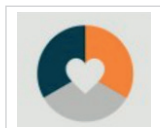
Legнды, w których wyjaśniane są dane często umieszczane są daleko od zawartości wykresu – z prawej strony lub pod wykresem. Zintegrowanie legendy z serią danych czyni dane bardziej dostępne dla odbiorcy.

PUBLICZNOŚĆ



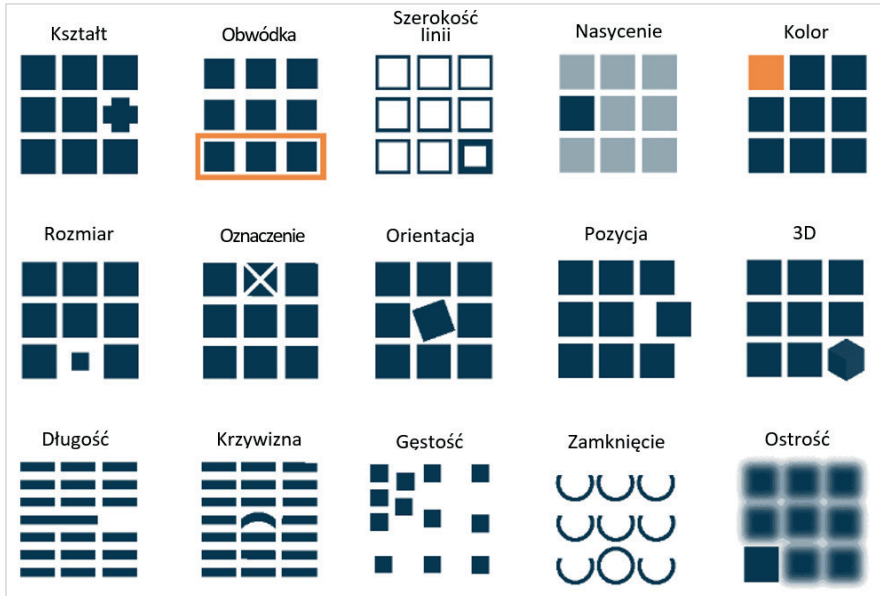
Zawsze należy brać pod uwagę odbiorców wizualizacji – czy potrzebują krótkiego raportu, czy może bardziej szczegółowego dokumentu, czy też interaktywnego narzędzia do eksploracji danych online.

Z ROZWAGĄ UŻYWAJ WYKRESÓW KOŁOWYCH



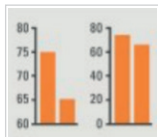
Ludzkie oko ma trudności w porównywaniu powierzchni. Często inne typy wykresów będą znacznie lepszym wyborem.

UWZGLĘDNIJ PRZETWARZANIE MIMOWOLNE



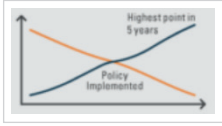
W skutecznej wizualizacji powinno się wykorzystywać przetwarzanie mimowolne (*preattentive processing*), podczas którego operacje poznawcze odbywają się przed skupieniem uwagi na określonym obszarze obrazu, czyli natychmiast po dotarciu sygnału z receptorów zmysłowych do mózgu. Oczy wykrywają ograniczony zestaw cech wizualnych, które łączymy i postrzegamy jako obraz, zanim staną się przedmiotem świadomego zainteresowania.

OŚ WYKRESÓW SŁUPKOWYCH I KOLUMNOWYCH ROZPOCZYNAJ W PUNKCIE ZERO



Wykresy słupkowe i kolumnowe, które nie rozpoczynają się w punkcie zero zniekształcają proporcje między wartościami. W przypadku niewielkich zmian ilościowych warto rozważyć wizualizację różnicy lub zmiany wartości.

DOŁĄCZ ADNOTACJĘ



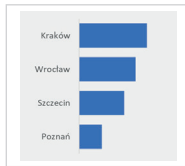
Tam gdzie to konieczne należy dodać tekst objaśniający, aby pomóc odbiorcy w odczytaniu i zrozumieniu wizualizacji oraz by prowadzić go przez jej treść.

UNIKAJ 3D



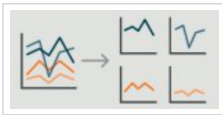
Korzystanie z wizualizacji 3D, gdy nie występuje trzeci wymiar (trzecia zmienna), zwykle utrudnia percepcję i zniekształca odbiór.

SPRAW, ABY ETYKIETY BYŁY ŁATWE DO ODCZYTANIA



W przypadku wykresów kolumnowych lub słupkowych warto zdecydować się na ten wariant, który pozwoli na umieszczenie etykiet w poziomie. Ogólnie należy dbać o to, aby etykiety były czytelne, zwarte i łatwe do zrozumienia dla odbiorcy.

W MIEJSCE ZŁOŻONEGO WYKRESU ZASTOSUJ KILKA MNIEJSZYCH

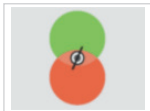


Podział skomplikowanego wykresu na odrębne mniejsze wizualizacje może być skutecznym sposobem na zwiększenie czytelności przekazu.

ZAGADNIENIA DOTYCZĄCE KOLORU I CZCIONKI



Unikaj domyślnych kolorów i czcionek – wszystkie wyglądają tak samo i nie wyróżniają się.



Rozważ ślepotę barw – około 10% ludzi (głównie mężczyzn) cierpi na jakąś formę ślepoty barw.



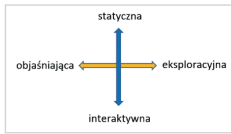
Unikaj palety kolorów tęczy – nie jest ona odwzorowana na nasz system liczbowy i nie ma logicznej kolejności.

OSTROŻNIE UŻYWAJ MAP



Zawsze należy upewnić się, że mapa jest poprawna, że takie miejsca/punkty geograficzne istnieją. Często wykresy kolumnowe lub słupkowe są lepsze do prezentacji porównań między jednostkami geograficznymi.

FORMA I FUNKCJA WIZUALIZACJI



Przed wyborem typu wizualizacji warto przeanalizować jeszcze dwa aspekty, tj. jej charakter (statyczny czy interaktywny) oraz funkcję (objaśnienie danych czy eksploracja).

Wskazówki dotyczące wizualizacji danych, w ramach upowszechniania dobrych praktyk, zamieściła też na jednej ze swoich stron (<https://www.geckoboard.com/best-practice/data-visualization-tips/>) firma programistyczna Geckoboard (<https://www.geckoboard.com/>), zajmująca się głównie tworzeniem i udostępnianiem narzędzi informatycznych do budowy pulpitów nawigacyjnych. Są one również do pobrania w formie plakatu, jako plik pdf (<https://www.geckoboard.com/uploads/play-your-charts-right.pdf>).

Autorzy podkreślają, że w świecie, w którym dane mają ogromne znaczenie (a w takim żyjemy obecnie), wszyscy chcą tworzyć efektywne wykresy. Ale ponieważ wizualizacji danych rzadko uczy się w szkołach lub podczas szkoleń w miejscu pracy, większość osób zdobywa niezbędne umiejętności w praktyce, tworząc grafikę i na bieżąco dokonując wyborów, tym samym popełniając błędy, które dezorientują odbiorców – od nadmiernego komplikowania lub przesadnego ozdabiania wykresu, po niepełne lub niedokładne przekazywanie treści, nieoddających znaczenia danych.

Z badań (Tal, Wanssink, 2016) wynika, że większość ludzi uwierzy we wszystko, co im się powie, o ile dane zobrazowane będą na wykresie. Biorąc zatem pod uwagę siłę przekonywania wykresów, ważne jest, aby właściwie z nich korzystać. Istnieje bowiem wiele różnych przyczyn, z powodu których wykres może przekazać niewłaściwą informację, zniekształcić fakty lub utrudnić albo wręcz uniemożliwić zrozumienie danych.

Poniżej, w skrócie omówiono kluczowe rekomendacje specjalistów z Geckoboard, w kolejności zgodnej z infografiką na plakacie (rys. 2.1).



Rysunek 2.1. Wskazówki dotyczące wizualizacji danych od firmy Geckoboard – plakat

Źródło: <https://www.geckoboard.com/uploads/play-your-charts-right.pdf>.

1. **Wykresy kolumnowe świetnie nadają się do porównań.** Są łatwe do odczytania, ponieważ „proszą” oczy odbiorcy o zrobienie prostej rzeczy – porównanie względnych wysokości kolumn. Aby jednak z takiego wykresu wyciągnąć sensowne wnioski, kolumny muszą być przedstawione w całości. A to oznacza, że oś pionowa musi zaczynać się w punkcie zero. Jeśli na wykresie widoczne są jedynie końce kolumn, to odbiorca traci możliwość dokonywania użytecznych porównań wizualnych. Jest wprowadzany w błąd i odbiera niewłaściwą wiadomość albo musi zwrócić uwagę na skalę i czytać liczby, co jest sprzeczne z ideą wykresu.
2. **Wykresy mają na celu udostępnienie danych szerokiemu gronu odbiorców. Należy przedstawić fakty bez nadmiernego komplikowania rzeczy.** Ludzie mają tendencję do kwestionowania swojej inteligencji i obwiniania siebie, gdy mają trudności z interpretacją wykresu. W większości przypadków jednak nie jest to wina odbiorców lecz wizualizacji. Jak pisze ekspertka w tym zakresie, Cole Nussbaumer Knaflitz (2015), stosunkowo łatwo zrobić coś prostego i później skomplikować to. Dużo trudniej jest zaprezentować coś złożonego w sposób przystępny dla odbiorców.
3. **Wykresy liniowe często pokazują trend.** Należy dbać o proporcje i unikać deformacji wykresu. Rozciąganie jego wysokości sztucznie wyolbrzymia prezentowaną sytuację, podczas gdy rozciąganie szerokości osłabia przekaz faktycznych danych.
4. **Wykresy kołowe wydają się przyjazne, ale w rzeczywistości są trudne do odczytania.** W większości przypadków można znaleźć lepszą alternatywę. Trudność w interpretacji wynika z tego, że mózg ludzki nie jest w stanie dokonać dokładnych szacunków ani porównań kątów. Jeśli części koła są zbliżone rozmiarem, odbiorcy trudno stwierdzić, która z nich jest większa, a która mniejsza i musi polegać na czytaniu opisów i etykiet. Trudniej porównywać powierzchnie lub objętości niż długości. Aby uniknąć złego wrażenia, należy używać ich w ostateczności.
5. **Gdy odbiorca musi ocenić i porównać dwa wymiary naraz, na przykład dwa okręgi, rzadko kiedy jego szacunki są dokładne.** Zazwyczaj nie docenia wielkości większych okręgów i przecenia wielkość mniejszych. Dzieje się tak, ponieważ instynktownie ocenia długość i szerokość kształtów, a nie ich obszary. Podczas tworzenia lub czytania grafiki o proporcjonalnym kształcie warto pamiętać o prostej zasadzie – przy podwojeniu wysokości koła, jego powierzchnia rośnie czterokrotnie. Jeśli natomiast wybrany został wykres trójwymiarowy, okręgi nie reprezentują już obszaru, lecz objętość.

6. **W przypadku liczb dodatkowe miejsca po przecinku wyglądają imponująco i sugerują dokładność, ale często są bezcelowe.** Przesadna precyzja może niepotrzebnie przeciążać odbiorcę. Stephen Few (2012) pisze, że wybór odpowiedniego poziomu precyzji liczb sprowadza się do następującej praktyki – stopień dokładności nie powinien przekraczać poziomu potrzebnego do realizacji celów komunikacyjnych i potrzeb czytelników.
7. **Aby skupić uwagę odbiorcy na danych, należy unikać obramowań, cieniowania, tła, ciemnych linii siatki i niepotrzebnych etykiet.** Amerykański statystyk Edward Tufte (2007) w stosunku do tego typu dekoracji wnętrza grafiki użył terminu *chartjunk* (wszystkie elementy wizualne na wykresach, które nie są konieczne do zrozumienia przedstawionych informacji lub odwracają uwagę odbiorcy od tych informacji). Alberto Cairo na swojej stronie poświęconej wizualizacji (Cairo, 2018) wskazuje wręcz, że „biała przestrzeń nie jest pustą przestrzenią”.
8. **Jeśli nie przedstawiamy trzeciego wymiaru, nie używajmy formatu 3D** – wypacza on dane i utrudnia porównania. Większość ekspertów od wizualizacji danych zgadza się, że użycie efektu 3D jest w najlepszym razie niepotrzebnym, a w najgorszym dezorientującym i mylącym elementem. Zaburza zdolność odbiorcy do odczytywania wartości z akceptowalną dokładnością (Kirk, 2019). Na wykresie 3D tylna część wydaje się mniejsza niż przednia, a nawet jest za nią ukryta, co sprawia, że porównania są prawie niemożliwe. Odbiorca nie wie ponadto, który punkt na kształcie 3D mierzyć w stosunku do skali osi.
9. **Koloru należy używać do przekazywania informacji, a nie do dekoracji.** Zbyt wiele kolorów może mylić i wprowadzać w błąd. Podczas projektowania wykresu kolor może być zarówno przyjacielem, jak i wrogiem projektanta. W zależności od tego, jak go użyje, może albo umiejętnie podkreślić dane i pokazać zmianę, albo stworzyć wizualne przeciążenie i zdezorientować odbiorcę. Ekspert od wizualizacji danych Stephen Few (2013) wskazuje, że jeśli jakkolwiek kolor wydaje się kontrastować z normą, odbiorca zwróci na to szczególną uwagę i postara się przypisać znaczenie tej różnicy. Zbyt wiele znaczeń lub kolorów zaciera przekaz i zniechęca ludzi.
10. **Częstym błędem, który popełniają projektanci wykresów, jest przesadne stosowanie zbędnych elementów, takich jak nadmiar koloru, zbyt wiele miejsc po przecinku, efekty specjalne, które tworzą bałagan graficzny.** Należy tego unikać, bo takie szczegóły nie zrobią na nikim wrażenia, a tylko utrudniają przekaz. Zgodnie ze słowami francuskiego pisarza Antoine de Saint-Exupéry

„doskonałość osiąga się nie wtedy, gdy nie ma nic więcej do dodania, ale wtedy gdy nie można już nic odjąć”.

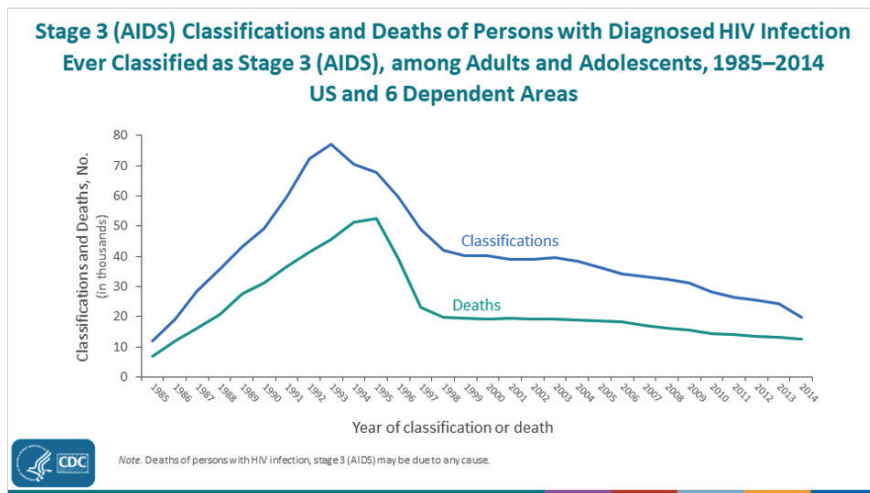
11. **Należy unikać niedopowiedzeń. Tekst jest przyjacielem projektanta.** Przemyslane użycie etykiet pomoże odbiorcy w szybkim przyswojeniu informacji. Odpowiedni opis (tytuł etykiety) to najprostszy, ale często najbardziej zaniebdywany krok przy udostępnianiu wykresów. Dodanie odpowiedniego tekstu, który jest zwięzły i trafny, pomaga odbiorcy wykorzystać potencjał umysłowy do zrozumienia danych. Cole Nussbaumer Knaflitz (2015) wskazuje, że tekst odgrywa wiele ról w przekazywaniu informacji: może służyć do oznaczania, przedstawiania, wyjaśniania, wzmacniania, podkreślania, polecania i opowiadania historii. Podstawowe teksty, jakie należy rozważyć dotyczą tytułu wykresu, tytułu osi, etykiet osi i etykiet danych. Często pomijane są jednostki miary i źródła danych, a one również odgrywają ważną rolę ułatwiającą interpretację wykresu. Cały czas należy pamiętać, że słowa są po to, aby pomóc, a nie odciągać uwagę od danych.
12. **Wykresy dwóch zestawów danych z jedną skalą po lewej i drugą po prawej mogą być mylące i sugerować związek, którego często nie ma.** Zagadnienie to omawia Stephen Few (2012), który wskazuje, że stosunkowo łatwe i czasami pomocne jest wykreślenie wielu zestawów danych przy użyciu wspólnej osi poziomej, jeśli wszystkie dane są wyrażone w tej samej jednostce miary. Ale jeśli jednostki są różne, należy dodać pomocniczą oś pionową po prawej stronie wykresu. Taki sposób prezentacji wymaga czasu i wysiłku ze strony odbiorców, aby rozszyfrować i zrozumieć, które dane należy odczytać względem której osi. Nawet jednak gdy rozwiążą tę zagadkę, kusi ich, aby porównywać wartości obu zestawów danych, co nie ma sensu, biorąc pod uwagę, że skale i jednostki danych są różne. Zasadniczo zatem należy unikać prezentacji zbyt dużej liczby danych na małej przestrzeni. Lepiej w takich przypadkach użyć dwóch lub więcej wykresów.
13. **Prezentacja danych w tabelach może być wygodna, gdy trzeba wyszukać poszczególne liczby. Aby taka tabela spełniała swoje zadanie, musi pokazywać dane w sposób ułatwiający znajdowanie liczb i porównywanie wierszy.** Wycenowanie liczb lub wyrównanie ich do lewej strony w komórkach niweczy cel i zmusza ludzi do „przeskakiwania” tam i z powrotem między wartościami, co utrudnia porównywanie wielkości. Wyrównanie liczb do prawej jest wskazane zarówno dla wartości całkowitych, jak i dziesiętnych. W przypadku wartości z cyframi dziesiętnymi Stephen Few (2012) sugeruje wyrównanie zarówno przecinka, jak i ostatniej cyfry do prawej. Można to osiągnąć, wyrażając każdą

wartość przy użyciu tej samej liczby cyfr dziesiętnych, nawet jeśli są to zera, ponadto dodanie separatorów tysięcy – takich jak np. kropki lub spacje – w celu podzielenia liczb na mniejsze części, może pomóc ludziom łatwo wykrywać różnice między liczbami.

14. **Prostsze jest zawsze lepsze.** Projektanci często zadają sobie pytanie – jakiego wykresu użyć? Większość ekspertów od wizualizacji danych zgadza się, że zawsze lepiej jest używać form graficznych, które ludzie znają.
15. **Słupki, linie, tabele lub, w większości przypadków, po prostu liczby to najskuteczniejsze sposoby przedstawiania danych przy minimalnym nakładzie pracy i zajmowanej przestrzeni.** Cole Nussbaumer Knaflitz (2015) zwraca uwagę, że to, iż mamy jakieś liczby, nie oznacza, że potrzebujemy wykresu, aby je pokazać. Kiedy próbujemy przedstawić tylko jedną lub dwie liczby na wykresie, medium odciąga uwagę od danych. Bezpośrednie pokazywanie liczby jest znacznie prostsze i ma większy wpływ.

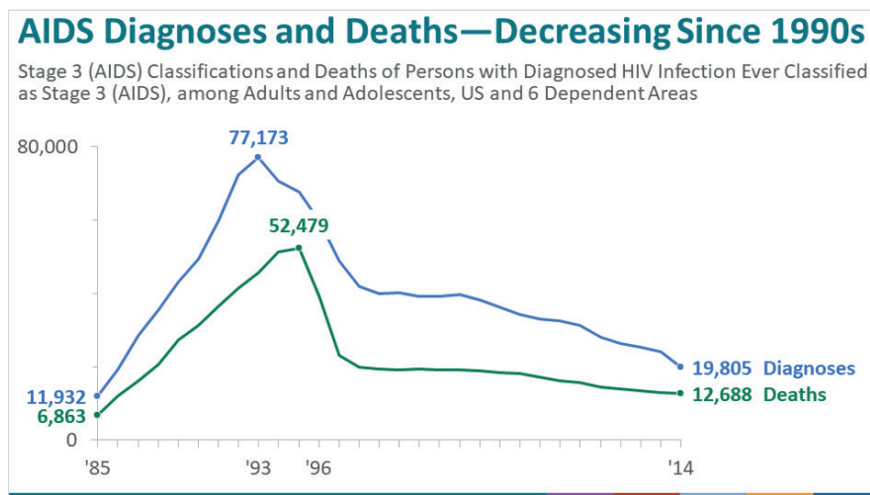
Cenne rekomendacje dotyczące tworzenia skutecznych wizualizacji znaleźć można również na witrynie Depict Data Studio (<https://depictdatastudio.com>), prowadzonej przez Ann K. Emery – cieszącą się międzynarodowym uznaniem specjalistkę od przygotowywania raportów i prezentacji. Rad udziela głównie podczas licznych warsztatów, webinarów i szkoleń dla firm, organizacji oraz wszystkich innych zainteresowanych efektywną wizualizacją danych. Są one również dostępne na blogu (<https://depictdatastudio.com/blog/>), a w szczególności we wpisach poświęconych wizualizacji danych (<https://depictdatastudio.com/category/data-visualization/>). Kwestie związane z przejrzystością i dostępnością wizualizacji Emery zebrała w e-booku (Emery, 2021). Omawia w nim aspekty kluczowe dla czytelników i przystępnych prezentacji danych. Na przykładach Emery przedstawia zagadnienia związane ze stosowaniem kolorów na wykresach, uzasadnia dlaczego należy unikać zestawienia zielonego i czerwonego, porusza też kwestie kontrastu, wielkich i małych liter, wskazuje na zalety umieszczania tekstów w poziomie, podpowiada, jak i kiedy stosować legendy i etykiety danych. Przedstawia również propozycję wyrównania elementów wizualizacji (teksty do lewej, liczby do prawej) oraz stosowania ikon tematycznie związanych z danymi (do pobrania np. tu: <https://thenounproject.com/>, <https://www.iconfinder.com/>), które to elementy mogą znacznie zwiększyć czytelność grafiki. Emery pokazuje również jak zredukować konieczność czytania i liczenia przez odbiorcę wizualizacji i przez to uczynić grafikę jeszcze bardziej dostępną. Jak uzasadnia, należy dążyć do obniżenia wysiłku, jaki odbiorcy muszą włożyć w czytanie i liczenie, nie dlatego, że są głupi, ale dlatego, że są zajęci. Muszą być w stanie zrozumieć przekaz za pierwszym razem, a nie po drugim,

trzecim czy czwartym przeczytaniu. Poniżej przedstawiono przykład wykresów przed i po poprawie czytelności, zaproponowanej przez Emery (rys. 2.2 i 2.3).



Rysunek 2.2. Wykres przed modyfikacjami zwiększającymi czytelność

Źródło: <https://depictdatastudio.com/accessibility-quick-wins-lower-the-reading-level/>.



Rysunek 2.3. Wykres po modyfikacjach zwiększających czytelność

Źródło: <https://depictdatastudio.com/accessibility-quick-wins-lower-the-reading-level/>.

Można zauważyć, że dokonano następujących zmian:

- wprowadzono zwięzły, prosty tytuł wykresu (jego czytelność w stosunku do wersji pierwotnej przetestowano na stronie <https://readable.com/>),
- pozostałe informacje zachowano w podtytule, przy czym zastosowano „hierarchię tekstu” i podtytuł sformatowano znacznie mniejszą i niepogrubioną czcionką,
- etykiety na osi X zapisano w formie skróconej, większą niż pierwotna czcionką, w układzie poziomym, przy czym podpisano jedynie kluczowe lata, w których obserwowane były wyraźne zmiany w danych,
- podobną operację zastosowano w stosunku do osi Y, etykiety na tej osi zapisano w rzeczywistych jednostkach, większą czcionką i również ograniczono je do trzech wartości kluczowych: minimalnych dla obu serii danych (sformatowanych kolorem odpowiadającym linii na wykresie) oraz wartości nieco większej niż maksymalna;
- zrezygnowano z tytułów osi.

Z kolei redukcję wysiłku związanego z liczeniem Emery przedstawia na następującym przykładzie (rys. 2.4):



Rysunek 2.4. Przekształcenia liczb, prowadzące do zmniejszenia wysiłku odbiorcy związanego z liczeniem i tym samym do zwiększenia czytelności wizualizacji

Źródło: <https://depictdatastudio.com/accessibility-quick-wins-lower-the-numeracy-level/>.

Pierwszym krokiem jest zaokrąglenie liczby do całkowitej, następnym etapem – przejście od zaokrąglonych wartości procentowych do ułamków. Ostatnim – zamiana ułamka na frakcję i jej użycie w zdaniu, np. 33,26% badanych to kobiety → 33% badanych stanowiły kobiety → 1/3 badanych stanowiły kobiety → jedna na trzy badane osoby była kobietą.

Ann K. Emery wspólnie z równie uznaną specjalistką od prezentacji danych – Stephanie Evergreen (<https://stephanieevergreen.com/about/>) opracowały tzw. listę kontrolną wizualizacji danych, będącą kompilacją 24 wytycznych, dotyczących poprawnego tworzenia i formatowania wykresów. Lista podzielona jest na 5 sekcji:

tekst, układ, kolor, linie i ogólne. Udostępniono ją na stronie: <https://stephanieevergreen.com/rate-your-visualization/>. Użytkownik może przesłać własną wizualizację i sprawdzić poprawność jej wykonania pod kątem spełnienia 24 zaleceń. Uzyskane wyniki może pobrać, by dopracować jakość przygotowanej grafiki, ale również może udostępnić je publicznie. Autorkom pozwala to rozwijać i kontynuować badania nad zasadnością listy kontrolnej, a dla innych użytkowników internetu stanowi dodatkowy materiał do analiz i inspiracji.

Omawiając zagadnienie dostępności, warto też wspomnieć, że w przypadku wizualizacji cyfrowych, zamieszczanych na stronach internetowych, blogach lub w mediach społecznościowych, istotne jest dodanie tekstu alternatywnego, zwanego także „atrybutem alt” lub „opisowym tekstem zastępczym”. Jest to tekst, który pojawi się zamiast grafiki lub zdjęcia, gdy z różnych powodów (np. słabego łącza) obraz na stronie nie może się wyświetlić. Dla osób z niepełnosprawnością wzroku jest najczęściej jedynym sposobem na zapoznanie się z przekazem graficznym. Według wytycznych WCAG 2.1 zapewnienie tekstów alternatywnych do każdej informacji nietekstowej jest obowiązkiem dla wszystkich instytucji publicznych. Podstawowe zasady tworzenia atrybutu alt są następujące:

- tekst powinien być krótki, ale wyczerpujący,
- należy pomijać wprowadzenie typu „na zdjęciu widać ...”, „grafika przedstawia ...” i od razu przejść do opisu zawartości,
- tekst powinien zawierać słowa kluczowe i być mocno osadzony w kontekście grafiki (tę samą grafikę można opisać na wiele różnych sposobów, w zależności do tego, na jakiej stronie się znajduje i jaki cel jej przyświeca),
- tekst powinien pełnić przede wszystkim funkcję użytkową (stąd należy unikać zbyt wielu epitetów, często pełniących funkcję oceniającą),
- należy unikać strony biernej, która nadaje tekstowi ton urzędowy i przeważanie utrudnia zrozumienie treści.

Krótki poradnik dotyczący tworzenia tekstów alternatywnych przygotowany przez A. Maciejewską z Fundacji Siódmy Zmysł, dostępny jest na witrynie „Otwarcia na różnice” Małopolskiego Instytutu Kultury <https://kulturawrazliwa.pl/wiedza/jak-stworzyc-tekst-alternatywny-krotki-poradnik/>.

Sporą porcję wiedzy w zakresie wizualizacji dostarcza również witryna Flowing-Data (<https://flowingdata.com/>) prowadzona przez Nathana Yau. Jest finansowana w 100% ze składek zarejestrowanych członków, którym w ramach opłaty abonamentowej Yau zapewnia m.in:

- samouczki „krok po kroku”, czyli setki godzin praktycznych instrukcji i wskazówek projektowych, jak tworzyć określone typy wykresów,

- kursy dotyczące wizualizacji danych,
- biuletyn przesyłany na maila, dostarczający praktycznych wskazówek na temat wizualizacji,
- dostęp do tysięcy postów i linków w archiwum FlowingData,
- dodatkowe zasoby w postaci linków i przewodników o sposobach tworzenia wizualizacji,
- porady eksperta, możliwość zadawania pytań i prowadzenia dyskusji dot. wizualizacji.

Warto też wspomnieć o projekcie Explore Explain (<https://www.visualising-data.com/podcast/>). Jest to seria dłuższych podcastów. Każdy odcinek oparty jest na rozmowie z projektantami grafiki, aby poznać historię stojącą za poszczególnymi wizualizacjami. Rozmowy są okazją do wyjaśnienia procesu projektowania i wglądu w niezliczone drobne decyzje, które leżą u podstaw ukończonych prac. Rzucają też światło na okoliczności, które ukształtowały myślenie projektanta. Widzowie mogą zrozumieć co, dlaczego i jak, poznając ukryte problemy i wyzwania, przełomy i objawienia, presję i frustrację i ostatecznie rzeczy, które zostały zrobione i te, z których zrezygnowano. Do tej pory ukazały się trzy sezony (w każdym po 7–9 odcinków), w których przeprowadzono rozmowy z utalentowanymi praktykami z całego świata, prezentując różne rodzaje prac wizualizacyjnych.

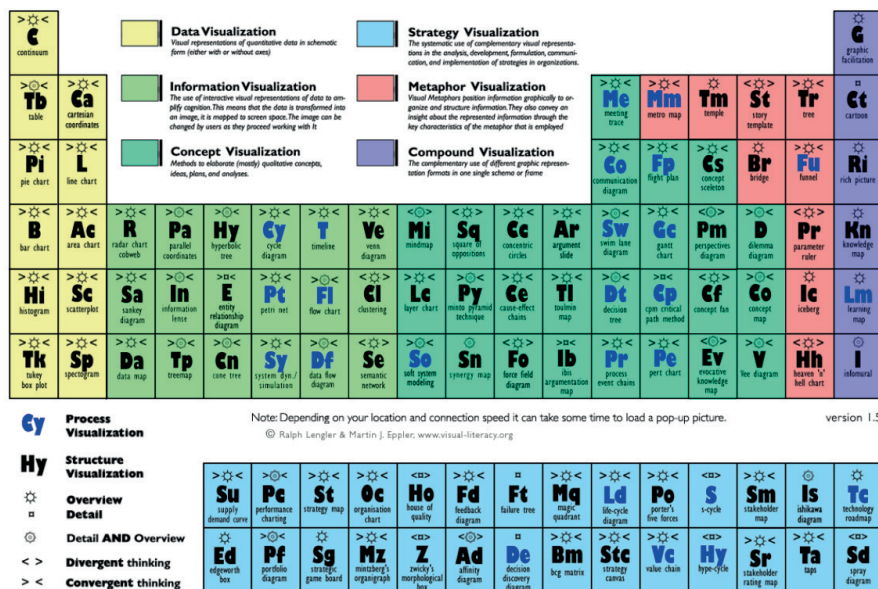
Oprócz podcastów na witrynie <https://www.visualisingdata.com/> prowadzonej przez Andy’ego Kirka dostępny jest blog (<https://visualisingdata.com/posts/>) z wieloma ciekawymi wpisami dotyczącymi tworzenia wizualizacji, jak również bogate zasoby (<https://www.visualisingdata.com/resources/>) z materiałami na temat przetwarzania danych, analityki, wykresów i aplikacji do tego służących.

2.2. Droga do wizualnej doskonałości

Wraz ze wzrostem zainteresowania wizualizacją, coraz bardziej potrzebne w celu uporządkowania wiedzy są taksonomie podejść wizualizacyjnych, ułatwiające korzystanie z wizualizacji w praktyce. Jedną z bardziej wszechstronnych tego typu taksonomii dostępna jest na witrynie projektu Visual Literacy (<https://www.visual-literacy.org/>), który koncentruje się na kluczowych (zwłaszcza dla ekspertów ds. komunikacji i analityków) umiejętnościach wizualnych, dotyczących tworzenia graficznej reprezentacji danych, informacji i wiedzy. Taksonomia zaprezentowana jest w formie „układu okresowego metod wizualizacji” (rys. 2.5). Pogrupowane są one według kategorii: wizualizacje danych, informacji, koncepcji, strategii, metafor i mieszane wizualizacje. Każda z nich jest następnie sklasyfikowana jako

wizualizacja procesu lub struktury i dalej zaszerzegowana w zależności od tego czy przedstawia szczegóły, przegląd, czy też szczegóły i przegląd oraz czy wspiera myślenie zbieżne czy rozbieżne.

W ramach tej taksonomii dostępnych jest ok. 100 typów wizualizacji. Po najechaniu kursorem myszy na określone pole wyświetla się przykład wizualizacji wybranego typu.



Rysunek 2.5. Układ okresowy metod wizualizacji opracowany w ramach projektu Visual Literacy

Źródło: https://www.visual-literacy.org/periodic_table/periodic_table.html.

Na stronie tego samego projektu (Visual Literacy) dostępna jest również grafika przedstawiająca tzw. schody do wizualnej doskonałości (rys. 2.6). Każdy stopień tych metaforycznych schodów podkreśla ważny wymiar projektu, jakim jest tworzenie graficznej reprezentacji danych, spostrzeżeń, doświadczeń lub wiedzy. Pierwszym krokiem jest dobranie odpowiedniego formatu wizualizacji i elementów graficznych. Drugi krok to podjęcie decyzji na temat sposobu wykorzystania takich cech obiektów graficznych jak: rozmiar, kształt, kolor, tekstura, kąt nachylenia czy pozycja, aby trafnie odwzorować przekaz. Trzeci krok dotyczy poszanowania

zasad projektowania przy wypełnianiu tła. Chodzi o stosowanie tych samych elementów dla tych samych funkcji oraz o konsekwencję w użyciu kolorów, kształtów i symboli, aby zaznaczyć jedność i ułatwić rozpoznanie. Podkreślenie zaś różnic może nastąpić poprzez użycie kontrastu. Umieszczanie podobnych elementów blisko siebie, używanie odpowiedniego wyrównania w celu grupowania i oddzielenia obiektów odmiennych sprzyja skupianiu uwagi odbiorcy oraz szybszemu wykrywaniu i zapamiętywaniu kluczowych elementów. Czwarty krok to sprawdzenie, czy udało się uniknąć potencjalnych pułapek wizualizacji, tj. czy nie została ona przeładowana i przeciążona danymi, czy nie nadużyto detali, czy nie zastosowano błędnie symboli lub kształtów, czy nie doszło do zniekształcenia danych lub do zbyt dużego uproszczenia komunikatu oraz czy wszelkie niejednoznaczności zostały wyjaśnione przez odpowiednie etykiety. Wreszcie krok piąty, sprawienie by wizualizacja była wyjątkowa, wnikliwa i atrakcyjna, by zaskakiwała i intrygowała odbiorców, wykraczała poza ich oczekiwania, przełamывała utarte schematy, zaciekała nowymi wzorcami metafor, pokazywała zagadnienie z nowej perspektywy.



Rysunek 2.6. „Schody do wizualnej doskonałości” – projekt Visual Literacy

Źródło: https://www.visual-literacy.org/stairs_of_viz/stairs_of_viz.html.

Nagrodą za dotarcie na szczyt schodów będzie wizualizacja przykuwająca uwagę odbiorców, spotykająca się z ich zrozumieniem, szybko zapadająca w pamięć i motywująca do przemyśleń.

Rolę spójności wizualizacji w procesie zrozumienia przekazywanych treści podkreślają też inni autorzy. Ware (2013) zaznacza, że przy wizualizacjach wykorzystywanych w praktyce zarządzania organizacjami, szczególnie dużymi korporacjami, niezbędne jest przestrzeganie pewnych standardów wizualizacji. Przekaz powinien być wolny od sprzeczności. Renaud i van Biljon (2017) wskazują, że każdy symbol użyty w wizualizacji powinien mieć jedno znaczenie lub reprezentować jedną koncepcję. To samo dotyczy kolorów i kształtów.

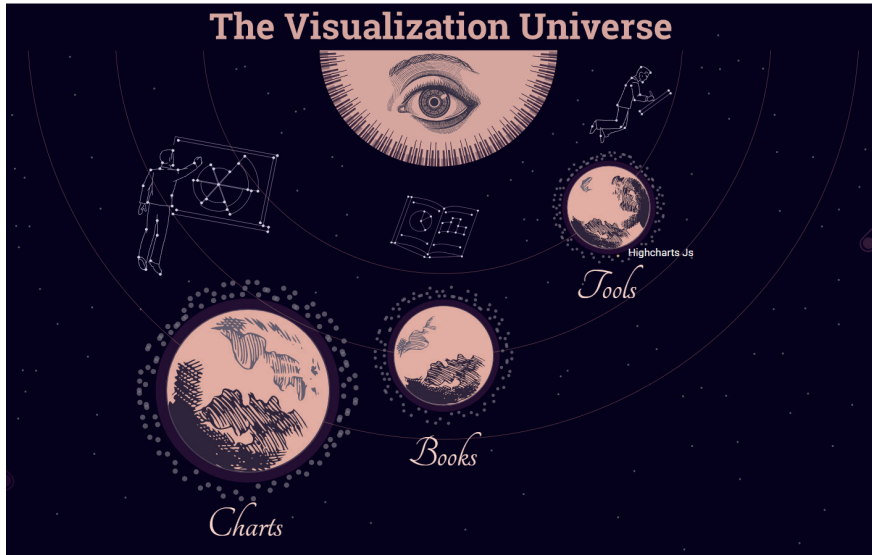
Aspekt jednoznaczności wizualizacji, która sprawia, że odbiorcy łatwo odnajdują znaczenie i rozumieją sens danych lub przekazu ponoszą również J.H. Mills, Thurlow i A.J. Mills (2010). Van Stralen (2015) wskazuje, że wątpliwości wymuszają poszukiwanie informacji, a wieloznaczność – poszukiwanie znaczenia. Bresciani i Eppler (2015) zwracają uwagę, że mimo, iż brak jednoznaczności może skłaniać do twórczej reinterpretacji wizualnego przedstawienia, a tym samym do nowych spostrzeżeń, to jednak może mieć również poważne negatywne skutki, gdy okaże się, że odbiorca nie jest w stanie dokonać interpretacji. Kryterium jednoznaczności można zatem uznać za dynamiczne i w dużym stopniu kontrolowane przez kontekst, w którym wizualizacja istnieje.

Z kolei Olshannikova i in. (2015) uzupełniają informacje w kwestii złożoności wizualizacji, o której decyduje zarówno jej forma, przedstawiona treść, jak i kontekst. Podkreślają, że zbyt wiele elementów w wizualizacji może spowodować, iż stanie się ona zbyt złożona, aby odbiorca mógł ją zrozumieć i aby mogła spełnić swoją rolę. Nie będzie wystarczająco przejrzysta i straci czytelność. Chodzi o sposób postrzegania, zauważania i rozróżniania elementów. To kryterium determinowane jest m.in. użytymi kształtami, ich układem, położeniem względem siebie, rozmiarem, kolorami, deseniami, kontrastem i efektami artystycznymi. Zbyt duże ich nagromadzenie może utrudniać zrozumienie wizualizacji. Układ i wygląd elementów powinien być taki, by wzmacniać uwagę, percepcję i ułatwiać interpretację wizualizacji (Marchese, Banissi, 2013).

Wartym uwagi źródłem wiedzy na temat wizualizacji, pozycji literaturowych z tego obszaru oraz dostępnych narzędzi jest interaktywny przewodnik o nazwie *The Visualization Universe* (wszechświat wizualizacji) <http://visualizationuniverse.com/>. Powstał w kooperacji portalu Adioma (<https://adioma.com/>), o którym jeszcze będzie mowa w tym rozdziale, oraz Google News Lab – zespołu w Google, którego misją jest „współpraca z dziennikarzami i przedsiębiorcami w celu budowania

przyszłości mediów”. Zasoby przewodnika (grafiki, książki i narzędzia do wizualizacji danych) zobrazowano symbolicznie w postaci kropek – księżyców krążących wokół swojej planety (rys. 2.7). Po kliknięciu myszką na wybraną kropkę pojawiają się informacje o danym elemencie, tj. krótki opis, link do strony internetowej ze szczegółowymi danymi, wykres przedstawiający popularność w ostatnim roku (wraz z linkiem do Google Trends – serwisu Google udostępniającego jeszcze więcej informacji na temat zapytań kierowanych do wyszukiwarki Google) oraz powiązane z nim inne tematy wyszukiwane przez internautów. Dane w projekcie podlegają codziennej automatycznej aktualizacji.

Intencją autorów projektu jest badanie: „Jakich rodzajów wizualizacji szukają użytkownicy i dlaczego?” oraz „Jakie narzędzia i książki wpływają na ich decyzje?”



Rysunek 2.7. Strona główna projektu „Wszechświat wizualizacji” prowadzonego przez portal Adioma i Google News Lab

Źródło: <http://visualizationuniverse.com/>.

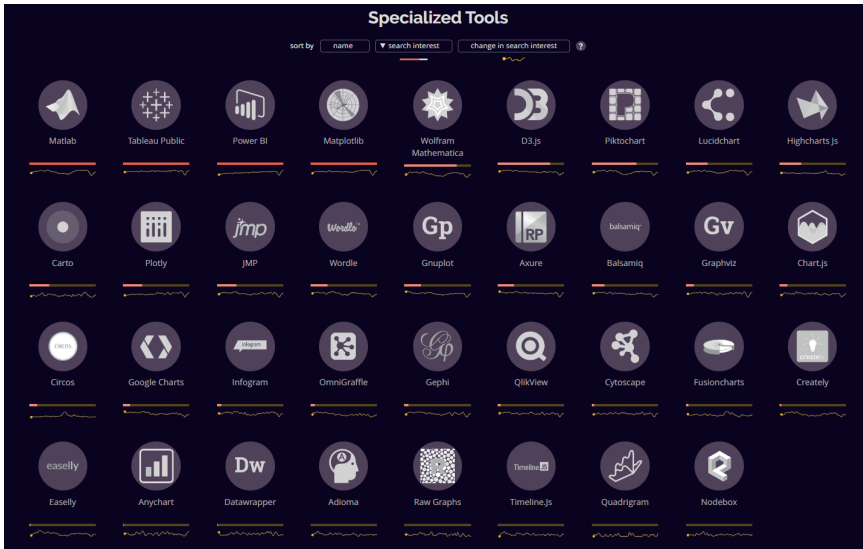
Wszystkie grafiki, książki i narzędzia służące do wizualizacji mogą być posortowane i wyświetlone na trzy sposoby:

- według średniego zainteresowania wyszukiwaniem w ciągu ostatniego roku,

- według zmiany zainteresowania wyszukiwaniem w ciągu ostatniego roku (w stosunku do roku poprzedniego),
- w porządku alfabetycznym.

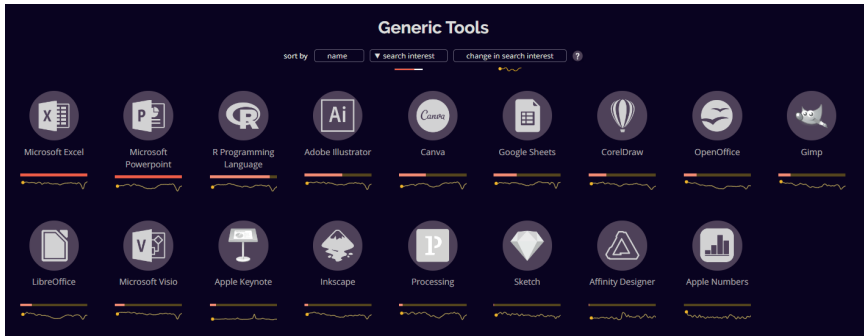
Kryterium doboru proponowanych w ramach projektu narzędzi była możliwość wygenerowania przez nie wizualizacji od początku do końca. Narzędzi, które nie spełniały tego kryterium, nie wzięto pod uwagę. Dodatkowo narzędzia podzielono na specjalistyczne, które służą wyłącznie do tworzenia infografik lub wizualizacji danych (rys. 2.8) oraz na ogólne, które oprócz tworzenia wizualizacji mają inne istotne funkcjonalności (rys. 2.9).

Autorzy zapraszają użytkowników do kontaktu i współpracy przy rozwijaniu platformy, do nadsyłania sugestii o dobrych książkach, wykresach i nowych narzędziach aktualnie nieuwzględnionych w projekcie.



Rysunek 2.8. Narzędzia specjalistyczne do wizualizacji danych analizowane w projekcie „Wszechświat wizualizacji” – sortowanie według zainteresowania wyszukiwaniem

Źródło: <http://visualizationuniverse.com/tools/>.

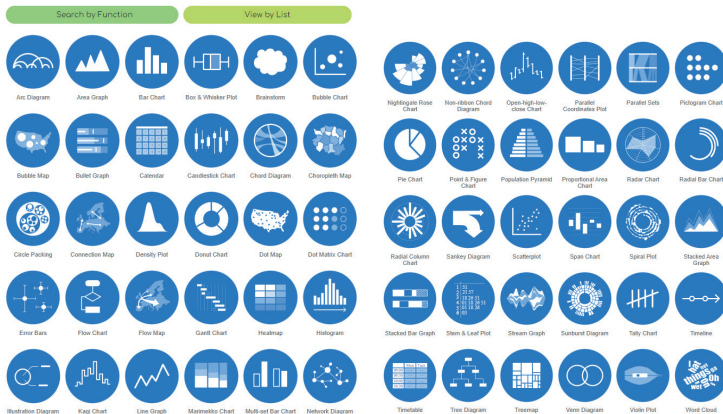


Rysunek 2.9. Narzędzia ogólne do wizualizacji danych analizowane w projekcie „Wszechświat wizualizacji” – sortowanie według zainteresowania wyszukiwaniem

Źródło: <http://visualizationuniverse.com/tools/>.

2.3. Katalog wizualizacji i selektory wykresów

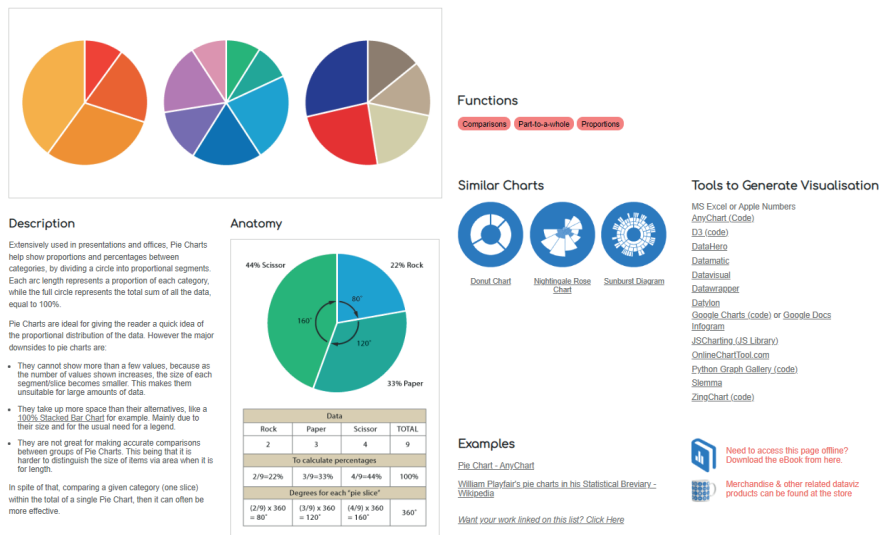
Severino Rebecca – niezależny projektant, konsultant i trener w zakresie wizualizacji, oferuje na swojej stronie internetowej (<https://datavizcatalogue.com/>) bezpłatny dostęp do katalogu wizualizacji danych (The Data Visualization Catalog). Zasób pełni rolę biblioteki referencyjnej różnych typów wizualizacji, które można filtrować według funkcji, aby wybrać odpowiedni typ wykresu dla swoich danych (rys. 2.10). Szczegóły każdej wizualizacji obejmują pełny opis i budowę, linki do przykładów oraz do narzędzi umożliwiających stworzenie wykresu.



Rysunek 2.10. Katalog wizualizacji danych opracowany przez Severino Ribecę

Źródło: <https://datavizcatalogue.com/>.

Po kliknięciu na wybrany typ wykresu można się zapoznać z jego szczegółowym opisem, budową, funkcjami, a także wadami i zaletami. Autor podaje też przykłady praktycznego zastosowania danego wykresu oraz wizualizacje, którymi można zastąpić omawiany typ. Wymienia również narzędzia informatyczne do generowania wybranej wizualizacji (rys. 2.11).



Rysunek 2.11. Szczegółowe informacje o wykresie kołowym – przykład z katalogu wizualizacji danych Severino Ribecci

Źródło: https://datavizcatalogue.com/methods/pie_chart.html.

Wybranie opcji wyszukiwania według funkcji powoduje przejście do podstrony z 16 kategoriami wizualizacji (rys. 2.12), które zawierają w sobie wiele różnorodnych form zobrazowania graficznego danych, pogrupowanych zgodnie z zastosowaniem lub tym, co wizualizacja ma zakomunikować odbiorcy. Autor klasyfikacji zastrzega przy tym, że nie jest ona perfekcyjna, jednak może być użyteczną wskazówką do wyboru grafiki odpowiednio do potrzeb komunikacyjnych. Do wyboru są takie kategorie jak:

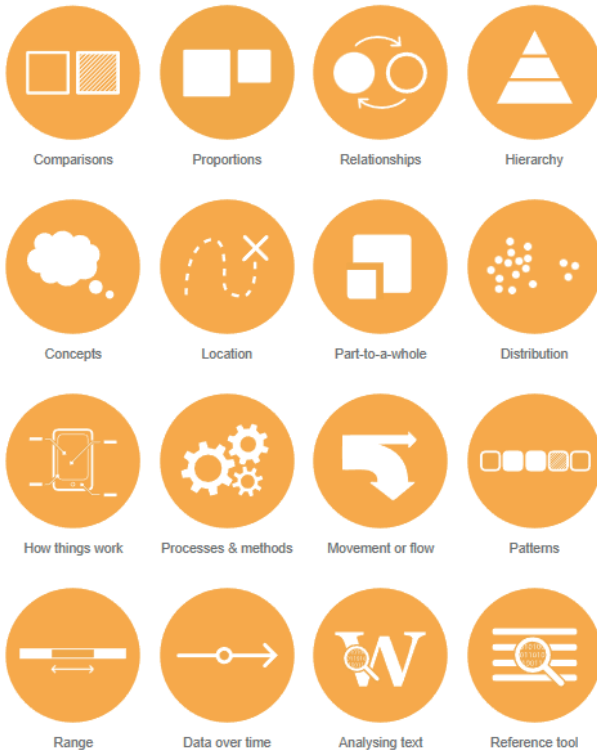
- porównania – wizualizacje, które pozwalają pokazać podobieństwa lub różnice między wartościami (np. wykres słupkowy, liniowy, warstwowy, bąbelkowy, kołowy, mapa cieplna, kartogram, wykres pudełkowy),

- proporcje – wizualizacje wykorzystujące powierzchnię do pokazania podobieństw lub różnic między wartościami lub stosunek części do całości (np. wykres bąbelkowy, mapa bąbelkowa, mapa drzewa, wykres pączkowy, kołowy, skumulowany wykres słupkowy, chmura wyrazowa),
- relacje – wizualizacje, które pokazują relacje i połączenia między danymi lub korelacje między dwiema lub większą liczbą zmiennych (np. wykres Marimekko, mapa cieplna, wykres współrzędnych równoległych, wykres radarowy, diagram Venna, schemat sieci, schemat drzewa, mapa połączeń, wykres bąbelkowy),
- hierarchie – wizualizacje, które pokazują, jak dane lub obiekty są uszeregowane i uporządkowane razem w organizacji lub systemie (np. schemat Sunburst, diagram drzewa, mapa drzewa, Circle Packing),
- koncepcje – wizualizacje, które pomagają wyjaśnić i zobrazować pewne idee (np. mapa myśli, diagram przepływu, diagram Venna, schemat ilustracyjny),
- lokalizacje – wizualizacje, które pokazują dane w regionach geograficznych (mapa bąbelkowa, kartogram, mapa połączeń, mapa punktowa, mapa przepływu),
- udział w całości – wizualizacje, które pokazują część (lub części) zmiennej do jej sumy, często są wykorzystywane do pokazania, w jaki sposób coś jest podzielone (np. wykres pączkowy, wykres Marimekko, wykres kołowy, skumulowany wykres słupkowy, schemat Sunburst, mapa drzewa),
- rozkłady – wizualizacje, które obrazują częstotliwość, rozmieszczenie danych w przedziale czasowym lub ich grupowanie (np. wykres gęstości, histogram, oś czasu, wykres pudełkowy, piramida populacji, wykres skrzypcowy, mapa połączeń, mapa punktowa, mapa przepływu),
- „jak to działa” – wizualizacje, które ilustrują sposób funkcjonowania obiektu lub system (np. wykres przepływu, schemat ilustracyjny, wykres Sankeya – strumieniowy),
- metody i procesy – wizualizacje pozwalające lepiej zrozumieć metodę lub proces (np. schemat blokowy, diagram Gantta, wykres Sankeya),
- ruch/przepływ – wizualizacje przydatne do pokazywania danych dotyczących ruchu lub przepływu danych (np. mapa połączeń, mapa przepływu, wykres Sankeya),
- szablony/wzorce – wizualizacje, które mogą pomóc ujawnić wzorce w danych i odkryć ich znaczenie (mapa kartogramów, punktowa, ciepła, połączeń, wykres bąbelkowy, radarowy, histogram, piramida populacji),
- zasięg/zakres – wizualizacje, które obrazują różnice między górnym i dolnym limitem na skali (wykres świecowy, pudełkowy, rozpiętości, Gantta, histogram),

- dane w czasie – wizualizacje, które pokazują dane w czasie, by uchwycić trend lub zmiany w czasie (np. wykres liniowy, świecowy, histogram, Gantt, oś czasu),
- analiza tekstu – wizualizacje, które ujawniają wzorce na podstawie analizy tekstu (np. chmura wyrazowa),
- narzędzia odniesienia – wizualizacje, które można wykorzystać jako narzędzie do łatwego wyszukiwania poszczególnych punktów danych (np. kalendarze, wykres Gantta, timetable, schemat drzewa).

What do you want to show?

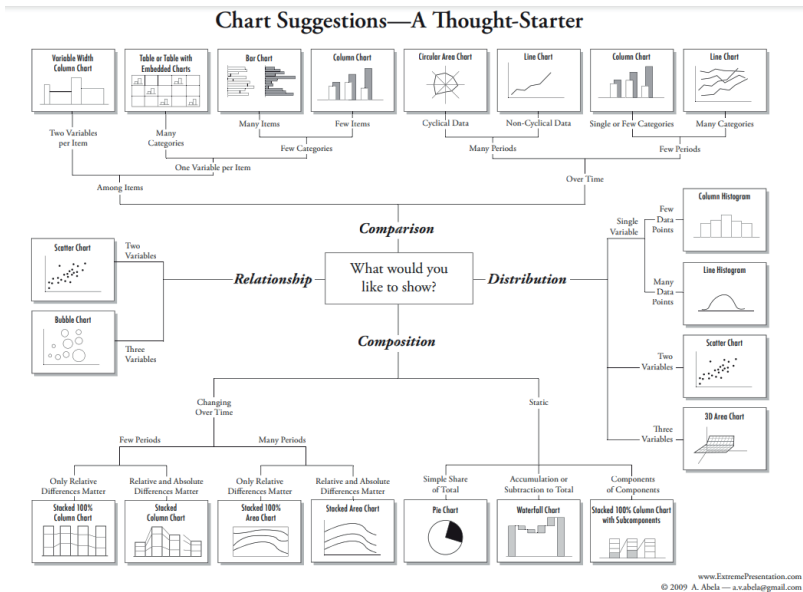
Here you can find a list of charts categorised by their data visualization functions or by what you want a chart to communicate to an audience. While the allocation of each chart into specific functions isn't a perfect system, it still works as a useful guide for selecting chart based on your analysis or communication needs.



Rysunek 2.12. Podział wykresów na kategorie w katalogu wizualizacji danych Severino Ribecci

Źródło: <https://datavizcatalogue.com/search.html>.

Pomoc przy wyborze wykresów zaproponował również Andrew Abela – specjalista od projektowania efektywnych prezentacji oraz autor metody Extreme Presentation™ (<https://extremepresentation.com/>). Abela opracował tzw. Chart Chooser, czyli przewodnik ułatwiający dobór wizualizacji w zależności od funkcji, a następnie od typu danych, jakie chce się zaprezentować (rys. 2.13). Należy zacząć od pytania postawionego na środku rysunku: „Co chcesz pokazać?”, a następnie przesuwając się na zewnątrz schematu, aby zidentyfikować typy wykresów, najbardziej odpowiednie do prezentacji danych.



Rysunek 2.13. Przewodnik po wyborze wykresów opracowany przez Andrew Abelę

Źródło: <https://extremepresentation.typepad.com/files/choosing-a-good-chart-09.pdf>.

Patrik Lundblad (<https://www.qlik.com/blog/patrik-lundblad/>) przygotował kolorową wersję przewodnika po wyborze wykresów opracowaną przez Abelę. Dodatkowo uzupełnił przewodnik o analizę każdej z czterech grup wykresów (<https://www.qlik.com/blog/third-pillar-of-mapping-data-to-visualizations-usage/>):

1. **Porównania** – te wizualizacje służą do porównania wartości ze sobą i mogą być używane do łatwego znajdowania wartości najmniejszych lub największych. Mogą być również stosowane do porównywania danych bieżących z wcześniejszymi, by sprawdzić, czy wartości rosną, czy maleją. Nie ma jednego idealnego typu wykresu, który można zawsze zastosować w celu porównania danych. W niektórych sytuacjach lepsze będą wykresy liniowe, w innych słupkowe lub warstwowe (dla danych cyklicznych). Przykładowe pytania, na jakie może odpowiadać taka wizualizacja to: „jakie produkty się najlepiej sprzedają” lub „jakie są wyniki sprzedaży w porównaniu z ubiegłym rokiem”.
2. **Zestawienia** – ten typ wizualizacji służy do zobrazowania, jak pewna część ma się do całości i jak można podzielić łączną wartość na udziały. Przeważnie wykresy pokazują wartość względną, ale niektóre mogą być również użyte do prezentacji różnicy bezwzględnej. Odnoszą się do zestawów danych, które zmieniają się w czasie lub zawierają dane statyczne. W przypadku danych statycznych sprawdzić się może wykres kołowy, choć istnieje wiele innych opcji, które równie dobrze go zastąpią. W przypadku danych, które zmieniają się w czasie, krytycznym zasobem jest liczba punktów danych. Należy również wziąć pod uwagę, że dane na osi muszą być przedstawione według czasu, a nie np. rosnąco według wartości. Przykładowe pytania, na jakie może odpowiadać taka wizualizacja to: „jak dużą część rynku mamy w danym regionie” lub „na jakie obszary został podzielony nasz budżet”.
3. **Rozkłady** – tego typu wykresy umożliwiają uzyskanie odpowiedzi na pytania o rozkład wartości liczbowych i częstotliwość ich występowania w zestawie danych oraz istnienie wartości odstających. Można je wykorzystać, aby zobrazić np. liczby klientów w poszczególnych grupach wiekowych lub liczebności osób, których zarobki należą do określonych przedziałów.
4. **Relacje** – tego rodzaju wykresy służą do wyświetlania relacji między danymi i mogą ułatwić znajdowanie korelacji lub wartości odstających. Przykładowe pytania, na jakie może odpowiadać taka wizualizacja to: „czy istnieje związek między wydatkami na reklamę a sprzedażą naszych produktów” lub „w jaki sposób wydatki i dochody różnią się w zależności od regionu i jakie są odchylenia”.

Selektor wykresów Abeli był inspiracją i podstawą do opracowania wersji online narzędzia, udostępnionego na stronie <https://www.juiceanalytics.com/chartchooser>. Zwiększono liczbę kategorii, według których można dokonać filtrowania: porównanie, rozkład, zestawienie, trend, relacja, tabela. Kolejne kroki to wybór programu (PowerPoint lub Excel), w którym chce się dokonać wizualizacji,

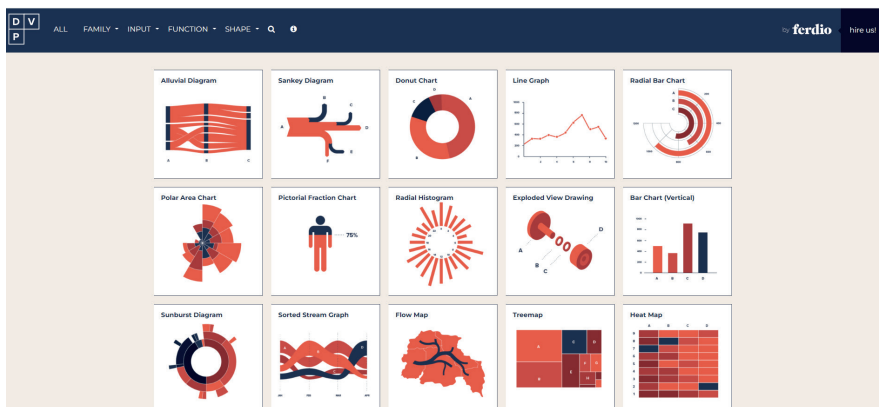
pobranie darmowego szablonu z wizualizacją i dostosowanie go do swoich danych (rys. 2.14).



Rysunek 2.14. Selektor wykresów online opracowany przez zespół z Juiceanalytics

Źródło: <https://www.juiceanalytics.com/chartchooser>.

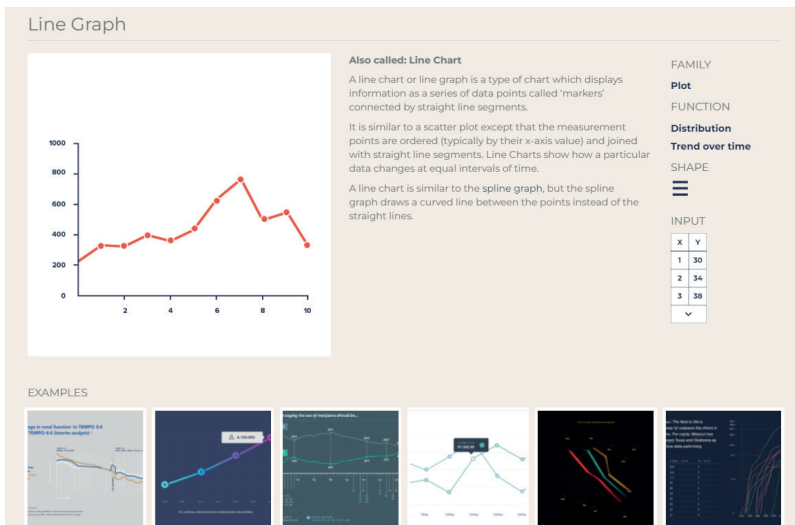
Kolejnym miejscem w internecie, które służy pomocą w zakresie doboru odpowiedniej metody wizualizacji jest strona projektu Data Viz (<https://datavizproject.com/>), prowadzona przez duńską agencję infografiki i wizualizacji danych Ferdio (<https://www.ferdio.com/>). Zaprezentowano i szczegółowo omówiono na tej stronie wiele istotnych rodzajów wizualizacji danych, które można filtrować według typu (wykres, diagram, mapa, graf, tabela), danych wejściowych (liczba i typ zmiennej), funkcji (porównania, wizualizacje koncepcji, korelacje, rozkłady, wizualizacje danych geograficznych, część do całości, trendy w czasie) i kształtu (obszar, słupki, koła, punkty, ikony, linie, mapy, wielokąty, piramidy, kwadraty). Obecnie witryna udostępnia ponad 160 wizualizacji (rys. 2.15).



Rysunek 2.15. Fragment witryny projektu Data Viz

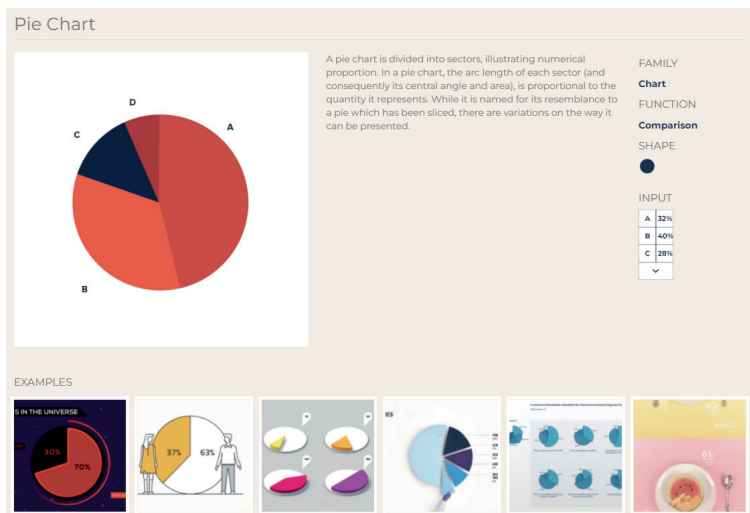
Źródło: <https://datavizproject.com/>.

Po najechaniu myszką i kliknięciu na wybrany typ wizualizacji pojawia się szczegółowy opis i linki do przykładowych wykresów danego typu (rys. 2.16, 2.17).



Rysunek 2.16. Fragment zrzutu ekranu dotyczący informacji o wykresie liniowym – projekt Data Viz

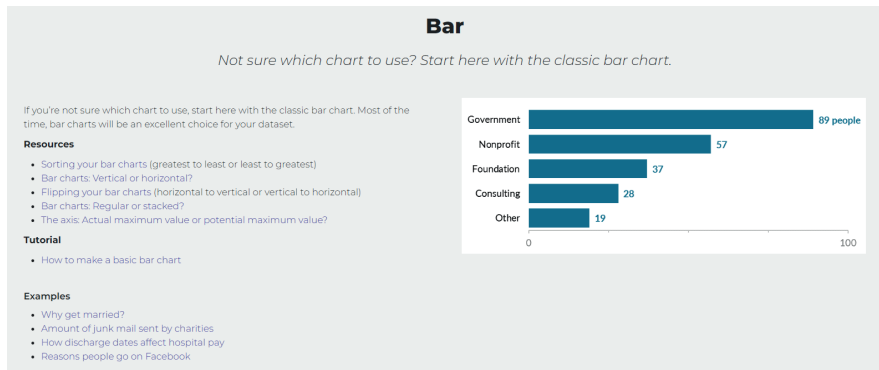
Źródło: <https://datavizproject.com/data-type/line-chart/>.



Rysunek 2.17. Fragment zrzutu ekranu dotyczący informacji o wykresie kołowym – projekt Data Viz

Źródło: <https://datavizproject.com/data-type/pie-chart/>.

Rozbudowany, interaktywny selektor wykresów dostępny jest też na wspomnianej już witrynie Depict Data Studio (<https://depictdatastudio.com/>), prowadzonej przez Ann K. Emery. Selektor pozwala na zastosowanie filtra dotyczącego liczby wizualizowanych zmiennych oraz funkcji wykresu. Każdy typ wykresu jest opisany przez podanie linków do zasobów dostępnych w ramach witryny Depict Data Studio oraz przykładów ze stron zewnętrznych. W przypadku prostszych wykresów jest też link do instruktażu, jak wykonać dany wykres. Tutorial zawiera zrzuty ekranu i prowadzi przez kolejne etapy tworzenia wizualizacji w MS Excel (rys. 2.18).



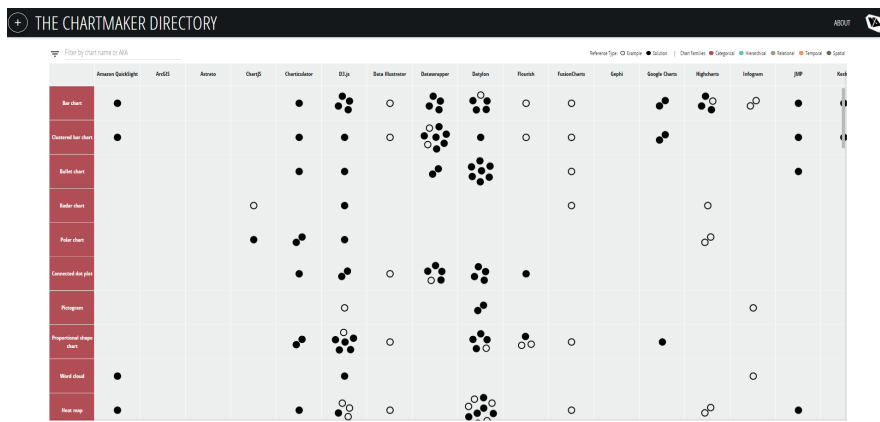
Rysunek 2.18. Fragment zrzutu ekranu dotyczący informacji o wykresie słupkowym z witryny Depict Data Studio

Źródło: <https://depictdatastudio.com/charts/bar/>.

Andy Kirk – badacz, szkoleniowiec, wykładowca, doradca i jednocześnie twórca oraz dyrektor zarządzający Visualizing Data (brytyjskiej spółki zajmującej się wizualizacją danych – <https://www.visualisingdata.com/>), opracował tzw. katalog kreatora wykresów (Chartmaker Directory), który jest dostępny na stronie <http://chartmaker.visualisingdata.com/>.

Witryna powstała w celu stworzenia przydatnego katalogu referencji, który pozwoli użytkownikom uzyskać odpowiedzi na często zadawane pytania, typu: „jeśli chcę stworzyć określony wykres, za pomocą jakich narzędzi mogę to zrobić”, „jeśli chcę zainwestować w dane narzędzie, jakie wykresy będę mógł/mogła wykonać”, „gdzie znajdę wskazówki, jak stworzyć konkretny wykres w danym narzędziu”, „jestem dostawcą, który chce zwiększyć możliwości mojego narzędzia i chce się dowiedzieć, jakich typów wykresów brakuje i które wydają się być niedostatecznie obsługiwane przez oferty konkurencji”. Aby ułatwić znalezienie odpowiedzi na

te i podobne pytania, zawartość katalogu zaprezentowana jest w postaci matrycy. W kolumnach wyszczególniono narzędzia do tworzenia wykresów, a w wierszach znajduje się obszerna lista różnych typów wykresów. W przecinających się komórkach umieszczono znaczniki (kropki), reprezentujące odniesienie do przykładu (konkretnego wykresu wykonanego w danym narzędziu – biała kropka) lub do rozwiązania (np. wskazówek w postaci samouczków z instrukcjami krok po kroku, prezentacji wideo, szablonów do pobrania – czarna kropka). Po najechaniu myszką na dany wykres pojawia się ponadto krótka informacja o nim oraz inne nazwy, pod którymi może być znany (rys. 2.19).



Rysunek 2.19. Fragment zrzutu ekranu z katalogiem kreatora wykresów (Chartmaker Directory), opracowanym przez Andy'ego Kirka

Źródło: <http://chartmaker.visualisingdata.com/>.

Kirk wyraźnie zaznacza, że rozległe luki w danym narzędziu (tj. brak odniesień) nie muszą oznaczać, że jest ono gorsze od innych. Jest bowiem wiele narzędzi bardzo specjalistycznych, których celem nie jest zaoferowanie użytkownikom możliwości tworzenia różnorodnych wykresów, lecz zapewnienie bogatej funkcjonalności w pewnym obszarze, jak na przykład Mapbox, który skupia się na analizie i wizualizacji przestrzennej.

Jeśli chodzi o zestaw narzędzi, to Kirk ma świadomość, że kategoria ta szybko ewoluuje i w zasadzie nie ma możliwości zbudowania wyczerpującej listy. Nie zawsze też jest w stanie z całą pewnością stwierdzić, że dane narzędzie nie oferuje możliwości budowy danego typu wykresu, a dostawcy, niestaránnie prezentując

swoją ofertę, nie ułatwiają tego zadania. Kolejny problem polega na tym, że bardzo często funkcjonalności danego narzędzia są znacznie większe, gdy współpracuje ono z bibliotekami, pakietami rozszerzeń lub wtyczkami. Na przykład sam pakiet R może tworzyć wykresy, ale R z pakietem ggplot2 jest kombinacją, która odblokowuje możliwości kreowania znacznie bardziej wyrafinowanej wizualizacji. To, czy pakiet R powinien być prezentowany sam, czy z ggplot2 – i jak inne podobnie powiązane pakiety powinny być prezentowane – jest częścią dialogu, do którego Kirk zachęca internautów. Wszelkie sugestie dotyczące uzupełnień i udoskonalień są mile widziane przez autora. Pomoc w kształtowaniu witryny, przez włączanie do katalogu kolejnych narzędzi, sprawi, że będzie ona mogła dobrze pełnić swoją rolę.

W przeciwieństwie do narzędzi, zbiór wykresów nie jest katalogiem tak dynamicznym i szybko zmieniającym się. Kirk zawarł tu listę wykresów i własną taksonomię typów, o której szerzej pisze w swojej książce (Kirk, 2019). Wykresy są podzielone na pięć typów (kategoria, hierarchia, relacja, czas, przestrzeń) na podstawie celu analizy, którą mają ułatwiać. Typy te nie są całkowicie rozłączne. Nad matrycą znajduje się pole, które umożliwia przeszukiwanie i filtrowanie listy typów wykresów za pomocą wprowadzonych słów kluczowych. W lewym górnym rogu znajduje się też przycisk +, który oferuje prosty trzyetapowy proces, pozwalający współtwórcom zasugerować link do przykładu lub wskazówek, jak utworzyć konkretny wykres w określonym narzędziu (rys. 2.20). Jak sam autor wskazuje, jednym z podstawowych celów tej witryny jest stworzenie możliwości do *crowdsourcingu* wartościowych referencji.

Rysunek 2.20. Fragment zrzutu ekranu pokazujący możliwość współtworzenia witryny z katalogiem kreatora wykresów, opracowanym przez Andy’ego Kirka

Źródło: <http://chartmaker.visualisingdata.com/>.

The screenshot shows a web interface titled "THE CHARTMAKER" with a close button (X) in the top left. Below the title is a section "Add to the directory" containing a three-step process:

- 1 Enter tool and chart type
- 2 Submit a link
- 3 Add details

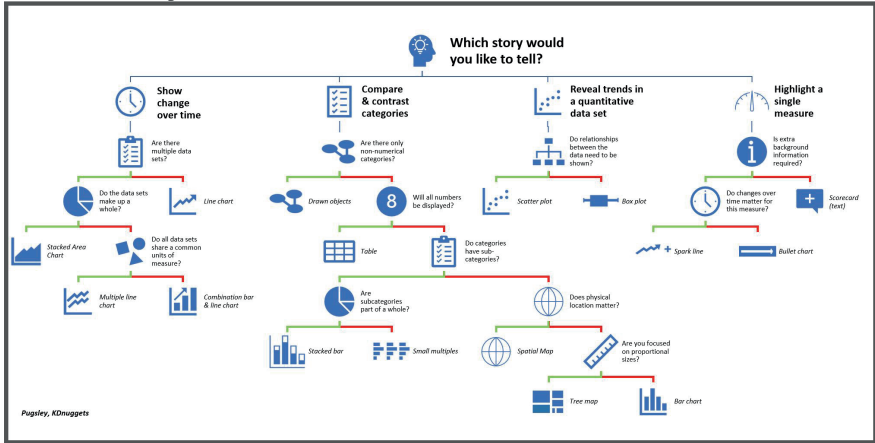
Below the steps, there are two input fields:

- TOOL**: Click on a cell in the grid or enter manually
- CHART TYPE**: Click on a cell in the grid

At the bottom, there are two buttons: "BACK" and "NEXT".

Kolejnym narzędziem, pomocnym przy doborze wizualizacji, może być selektor wykresów w postaci drzewa decyzyjnego, zaproponowany przez Stana Pugsleya (2022) (rys. 2.21).

Data Story Visualization: A Decision Tree



Rysunek 2.21. Selektor rekordów w postaci drzewa decyzyjnego – propozycja Stana Pugsleya

Źródło: <https://www.kdnuggets.com/wp-content/uploads/pugsley-visualization-decision-tree-kd-v12.jpg>.

Własną propozycję selektora wykresów przedstawił też Steve Franconeri. Ideę swojego narzędzia wyjaśnił w artykule (Franconeri, 2019). Autor zwraca uwagę, że większość selektorów i skróconych przewodników po wizualizacjach, dostarczających bezpośrednich sugestii, jak reprezentować dane, oparte są na sklasyfikowaniu wykresów przede wszystkim według zadań dla odbiorcy, tj. czy projektant chce, aby odbiorca zobaczył ranking, zbadał rozkład, aby sprawdził relację, czy też aby dokonał porównania. Propozycja selektora Franconeriego bazuje z kolei przede wszystkim na strukturze danych (rys. 2.22).

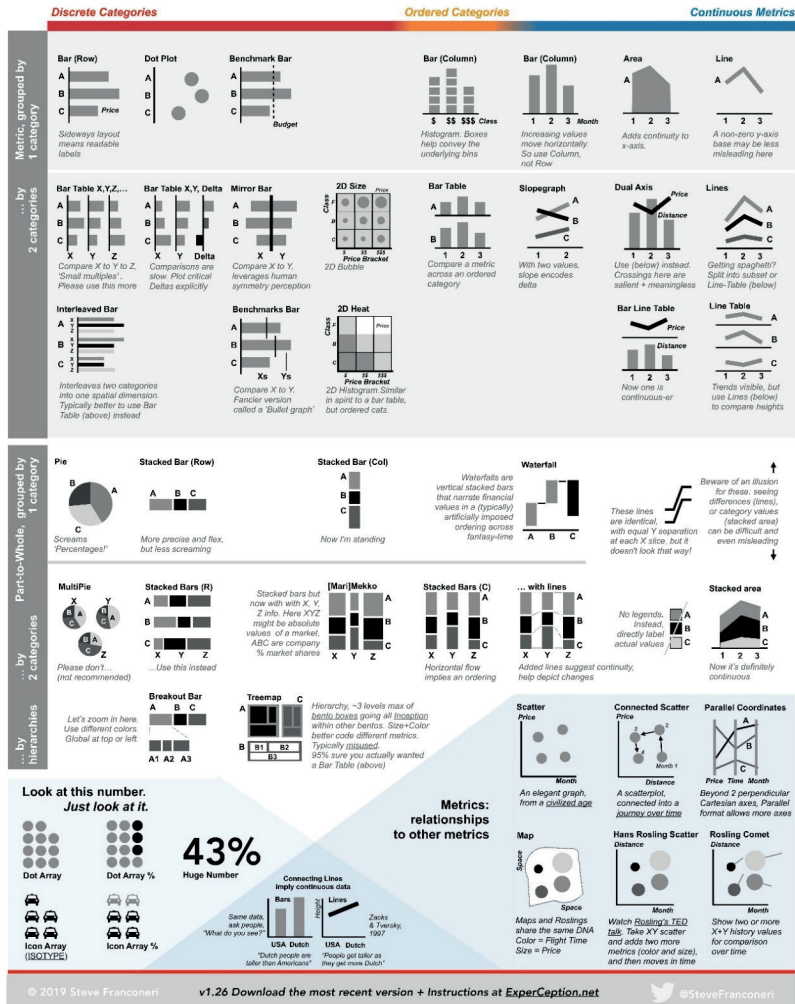
Which Visualization? A Quick Reference

You have the following data (sample):

Discrete Categories,
Ordered categories,
and Continuous Metrics

Here's how to plot them

Categories		Ordered Cats		Continuous Metrics			
City	Airline	Class	PriceBracket	Month	Distance	FlightTime	Price
Alphaville	XeroTrip	Coach	\$	1	300	120	250
Betastan	YoloFly	Business	\$\$	2	500	185	1,525
Chicago	ZeusAir	First	\$\$\$	3	650	240	4,023
***	***	***	***	***	***	***	***

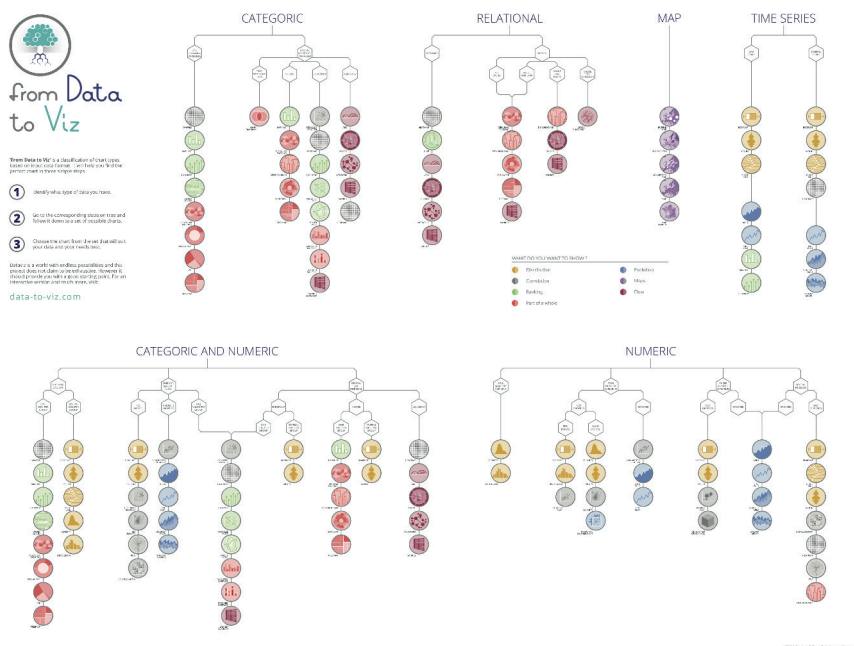


Rysunek 2.22. Jaki wykres wybrać – krótki poradnik Steve’a Franconieriego

Źródło: https://experception.net/Franconeri_ExperceptionDotNet_DataVisQuickRef.pdf.

Przy doborze odpowiedniego wykresu może również pomóc infografika przygotowana przez Stephanie Evergreen oraz karty „Chart chooser”, z których ideą można się zapoznać na stronie <http://chartchoosercards.com/>.

Doprowadzenie użytkownika do najbardziej odpowiedniego wykresu dla jego danych jest też celem projektu From Data to Viz (od danych do wizualizacji – <https://www.data-to-viz.com/>). Rozwiązanie bazuje na systematyzacji typów wykresów na podstawie formatu danych wejściowych. Klasyfikacja ma postać drzewa decyzyjnego prowadzącego do zbioru wizualizacji potencjalnie nadających się do zaprezentowania wejściowego zestawu danych (rys. 2.23). Dla poszczególnych typów wykresów podano przykład analizy opartej na rzeczywistych danych przy użyciu języka programowania R.



Rysunek 2.23. Drzewa decyzyjne ułatwiające dobór wykresu dla różnego rodzaju danych wejściowych

Źródło: <https://www.data-to-viz.com/poster.html>.

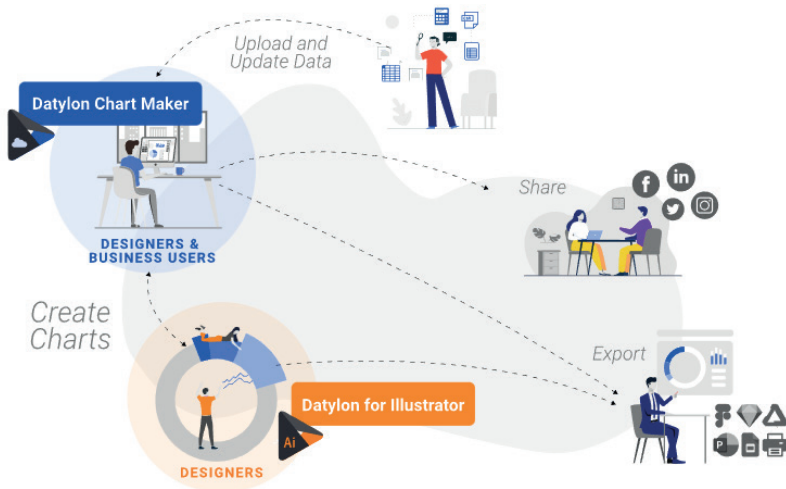
Po kliknięciu myszką na dany wykres pojawia się krótka informacja o nim wraz z odsyłaczem do znacznie pogłębionego opisu. Na końcu każdej „gałęzi” dostępny jest link „Story”, który przenosi użytkownika do kilku wizualizacji wykonanych na bazie rzeczywistych danych wraz ze sposobem ich tworzenia, omówieniem wad i zalet oraz potencjalnych problemów i elementów, na które trzeba zwrócić szczególną uwagę. W wielu miejscach omówienia są dodatkowe linki do stron poszerzających kluczowe pojęcia. Zestawy danych użyte w przykładach zamieszczone są w repozytorium udostępnionym użytkownikowi. Podano również kod w języku R do utworzenia grafiki.

Witryna dysponuje też zakładką <https://www.data-to-viz.com/caveats.html> zawierającą zbiór ostrzeżeń dotyczących najczęściej popełnianych błędów prowadzących do zniekształcania przekazywanych za pomocą wykresów informacji.

Jeśli chodzi o pomoc w zakresie konkretnego oprogramowania do wizualizacji, to należy też wspomnieć o witrynie <https://www.datawrapper.de/>, która promuje, i w ramach podstawowego pakietu udostępnia za darmo, narzędzie online do szybkiego tworzenia wykresów, map i tabel, które mogą być od razu osadzone na stronach www lub wyeksportowane do plików pdf i svg. Użytkownicy mają do dyspozycji blog z opisami najlepszych praktyk, materiały warsztatowe, ponad 100 artykułów wyjaśniających jak korzystać z aplikacji oraz przykładowe wizualizacje i szczegółowe analizy przypadków.

Warto też wymienić blog Gravy Anecdote (<https://gravyanecdote.com>), prowadzony przez Andy’ego Cotgreave’a – współautora wielkiej książki dashboardów biznesowych (Wexler, Shaffer, Cotgreave, 2017), wcześniej analityka Uniwersytetu Oksfordzkiego, a obecnie pracownika Tableau Software – amerykańskiej firmy informatycznej zajmującej się tworzeniem rodziny produktów do interaktywnej wizualizacji danych, specjalizującej się w analityce biznesowej. Wpisy na blogu dotyczą ogólnych tematów związanych z wizualizacją, ale też konkretnych samouczków dotyczących oprogramowania Tableau.

Narzędzie, które pomaga w tworzeniu wizualizacji i jej udostępnianiu oferuje też platforma Datylon <https://www.datylon.com/> (rys. 2.24). Rozwiązanie to nie wymaga żadnych umiejętności kodowania i jest przyjazne dla projektantów oraz użytkowników biznesowych. Do dyspozycji jest bogata baza szablonów <https://www.datylon.com/resources/chart-library> i przykładów <https://insights.datylon.com/stories>, którymi można się inspirować, projektując własny wykres. Proponowane wizualizacje można filtrować według typu wykresu.



Rysunek 2.24. Funkcjonalności oferowane na platformie <https://www.datylon.com/>

Źródło: <https://www.datylon.com/product>.

W ramach tego podrozdziału zdecydowano się również na bardziej szczegółowe przedstawienie wspomnianego już portalu Adioma (<https://adioma.com/>), prowadzonego przez Annę Vital, mimo że propozycja wykracza nieco poza tradycyjne metody wizualizacji danych liczbowych i obejmuje bardziej konceptualne style wizualne i diagramy często używane w środowiskach biznesowych. Autorka skupia się na tworzeniu infografik, czyli grafiki informacyjnej – wizualnej reprezentacji informacji, która wykorzystuje wzorce obrazów, aby opowiedzieć odbiorcom pewną historię i pomóc im zobaczyć, jak fragmenty informacji są ze sobą powiązane. Ważna jest obecność narracji dostosowanej do odbiorców, zgodnie z subiektywną wizją projektanta.

Portal Adioma, oprócz wielu ciekawych artykułów na temat infografiki, jej historii, znaczenia ikon przy projektowaniu, rodzajów i sposobów oddziaływania, oferuje też narzędzie do tworzenia infografiki o nazwie „Infographic Maker”. Wystarczy podać tekst, a system po przeanalizowaniu jego znaczenia zaproponuje grafikę i design. Można również samodzielnie, przeglądając liczne kategorie ikon (rys. 2.25), wybrać odpowiednie wizualizacje. Jest też opcja dodawania własnych obrazów. Szablony, które się same generują, dostosowują się do ilości tekstu.

Słowa można na bieżąco wprowadzać lub usuwać, a szablon adaptuje się do nowego układu (<https://adioma.com/features>).



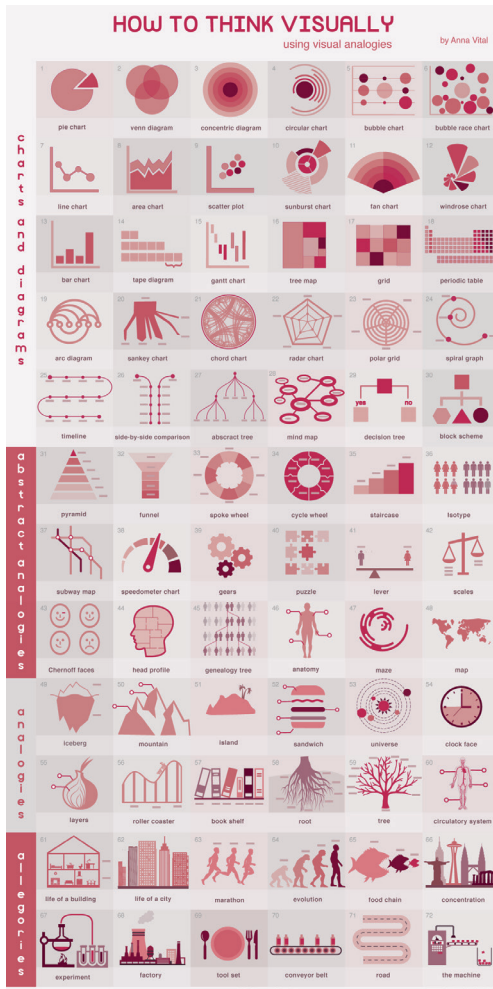
Rysunek 2.25. Kategorie ikon i przykładowe ikony w ramach kategorii, dostępne na portalu Adioma

Źródło: <https://adioma.com/icons>.

Infografiki na tym portalu budowane są na podstawie powszechnie rozpoznawalnych wzorców, takich jak oś czasu, siatka, drzewo czy cykl. W całej wizualizacji zachowany jest jeden styl graficzny ikon. Automatyczne schematy kolorów

zapewniają czytelność i harmonię odcieni. Dzięki „przełącznikowi szablonów” można spojrzeć na tę samą wizualizację wykonaną w różnych szablonach. Ostatecznie przygotowaną grafikę można łatwo wyeksportować do formatu wektorowego lub rastrowego, a także osadzić na stronie internetowej.

Swoje propozycje 72 różnych typów infografik Vital pogrupowała w cztery główne kategorie: wykresy i diagramy, abstrakcyjne analogie, analogie i alegorie (rys. 2.26). Każdą z wizualizacji opisano, do wielu podano przykład zastosowania (rys. 2.27).



Rysunek 2.26. Kategorie infografik zaproponowane przez Annę Vital

Źródło: <https://coolinfographics.com/blog/2015/9/1/how-to-think-visually.html>.

How to Make Your Writing Visual

by Anna Vital



Rysunek 2.27. Przykładowa infografika przedstawiająca, jak zwizualizować artykuł

Źródło: <https://adioma.com/@anna/infographic/how-to-make-an-article-visual>.

W podsumowaniu warto jeszcze raz przytoczyć wspomniany już blog Jona Schwabisha (<https://policyviz.com/>), gdzie w zakładkach (<https://policyviz.com/resources/> oraz <https://policyviz.com/books/>) znaleźć można wiele linków do blogów innych specjalistów od wizualizacji oraz do kluczowej literatury dla tej tematyki (Cairo, 2012; 2016; 2020; Camões, 2016; Few, 2009; 2012; Kirk, 2019; Meirelles, 2013; Nussbaumer Knaflic, 2015; Riche i in., 2018; Robbins, 2013; Schwabish, 2016; 2020; Wong Dona, 2013; Yau, 2011), a także do stron z oprogramowaniem do wizualizacji, takim jak: D3.js (biblioteka JavaScript do tworzenia dynamicznych, interaktywnych wizualizacji danych w przeglądarkach internetowych),

Datawrapper, Excel, Flourish, HighCharts, Infogram, JMP (interaktywna biblioteka do tworzenia wykresów i analizy danych połączona z SAS), Plotly, PowerBI, Processing, RAW Graphs, Qlik, Quadrigram, pakiet R, Tableau, Vega, Vizzlo.

Jeszcze jedna witryna jest godna polecenia (<https://coolinfographics.com/>), gdzie, jak pisze jej twórca Randy Krum – założyciel i prezes firmy zajmującej się wizualizacją danych i projektowaniem infografik InfoNewt (<https://infonewt.com/>) – znaleźć można wszystko, co związane jest z wizualizacją danych, infografikami, prezentacjami, pulpitami nawigacyjnymi i komunikacją wizualną. Podstrony witryny są kopalnią inspiracji, zawierają dziesiątki linków do portali z selektorami wykresów, przewodnikami referencyjnymi, ściągawkami, zbiorami narzędzi do projektowania grafiki oraz artykułami dotyczącymi wizualizacji.

3. Manipulacje i przekłamania na wykresach

3.1. Podatność wykresów na zniekształcenia

Jak już sygnalizowano w rozdziale pierwszym, przekaz wizualny ma znacznie większą siłę perswazji niż zwykły tekst, bowiem mocno wpływa na bodźce wzrokowe. Dane lub informacje zaprezentowane na wykresie lub prostej infografice są znacznie lepiej dostrzegalne i silniej zapadają w pamięć odbiorcy. Twórca wizualizacji ma bardzo duży wpływ na sposób jej percepcji przez adresata, co bywa świadomie wykorzystywane w celu wprowadzenia odbiorcy w błąd i doprowadzenia go do fałszywych wniosków, sprzecznych z faktami. Mimo że przeważnie wizualizacja opatrzona jest odpowiednim tekstem wyjaśniającym szczegóły i zależności, to jednak odbiorca przede wszystkim zwróci uwagę na obraz i na jego podstawie może wyrobić sobie zdanie na dany temat, rezygnując już z czytania opisu.

Najbardziej podatne na zniekształcenie i zabiegi manipulacyjne są wykresy i grafy, choć dezorientację odbiorcy można wywołać przez zmianę samej wielkości i rodzaju czcionki w tekście. Zgodnie z teorią Daniela Kahnemana, dotyczącą łatwości poznawczej, odbiorca komunikatu podświadomie robi wszystko, by ograniczyć wysiłek poznawczy, dlatego ważne jest, by tworząc przekaz, zadbać o jego maksymalną czytelność. Noblista (Kahneman, 2012) podaje przykład dwóch zdań, w których zapisano datę urodzin Hitlera (w obu przypadkach błędną). W pierwszym zdaniu czcionka była pogrubiona i większa od tej użytej w drugim zdaniu. Okazuje się, że większość osób tę pierwszą datę uznawała za poprawną. Egzemplifikacja ta pokazuje, że graficzne szczegóły wizualizacji wpływają na odbiór informacji. W przypadku komunikatów drukowanych podobnie oddziałuje lepszej jakości papier, większy kontrast między literami a tłem, czysty żywy kolor zamiast barw stonowanych. Odbiorca będzie miał większe zaufanie do prawdziwości komunikatów podanych w takiej formie.

W niniejszym rozdziale zwrócono uwagę na możliwości wypaczania faktów i manipulacji przy przekazie informacji, jakie mogą zastosować projektanci

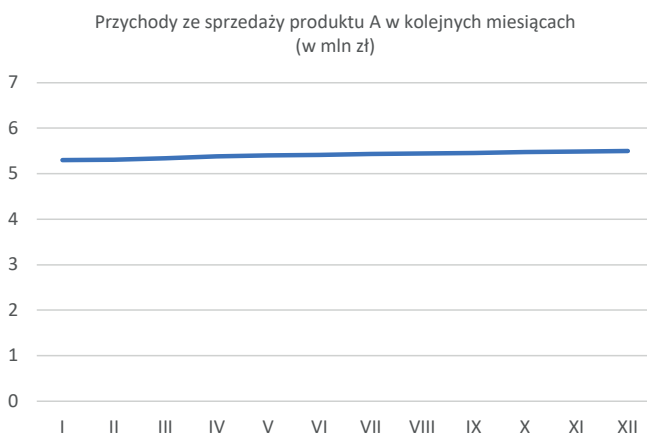
wykresów. Na fikcyjnych przykładach przeanalizowano przekłamania, wynikające ze świadomego wykorzystania różnych funkcji arkuszy kalkulacyjnych – jako jednych z najbardziej popularnych programów do tworzenia wykresów – celem przekazania treści zaplanowanej przez projektanta grafiki. Zaprezentowano też umyślne zniekształcenia informacji, wynikające z uwzględnienia na wykresach wybiórczych danych lub celowego doboru formy prezentacji „ukrywającej” rzeczywiste dane.

Świadomość tych zagadnień powinna wpływać na zachowanie ostrożności podczas odczytu prezentowanych informacji. Jest to szczególnie ważne w przypadku kadry menedżerskiej, dla której zestawienia i raporty bogate w wizualizacje w postaci wykresów, są często podstawą podejmowanych działań biznesowych.

W kolejnych podpunktach omówiono i zobrazowano różne rodzaje manipulacji na wykresach. Większość z omówionych przykładów była zaprezentowana przez autorkę książki podczas wystąpienia na 25. międzynarodowej konferencji „Knowledge-Based and Intelligent Information & Engineering Systems” (Krok, 2021).

3.2. Zmiana skali

Na wykresie 3.1 przedstawiono przychody ze sprzedaży produktu A w kolejnych miesiącach poprzedniego roku. Widać, że są one stabilne i mimo że ich przyrost jest bardzo powolny, to utrzymują się na w miarę stałym poziomie. Produkt ma zatem ugruntowaną pozycję, choć rynek jest już nasycony i być może w najbliższym czasie rentowność towaru będzie malała.



Wykres 3.1. Przychody ze sprzedaży

Źródło: opracowanie własne.

Jeśli jednak dział marketingu firmy sprzedającej produkt A chce, aby odbiorca otrzymał informację, że z miesiąca na miesiąc produkt A cieszy się coraz większym zainteresowaniem klientów, a jego sprzedaż dynamicznie rośnie, wystarczy zmienić skalę (wykres 3.2).

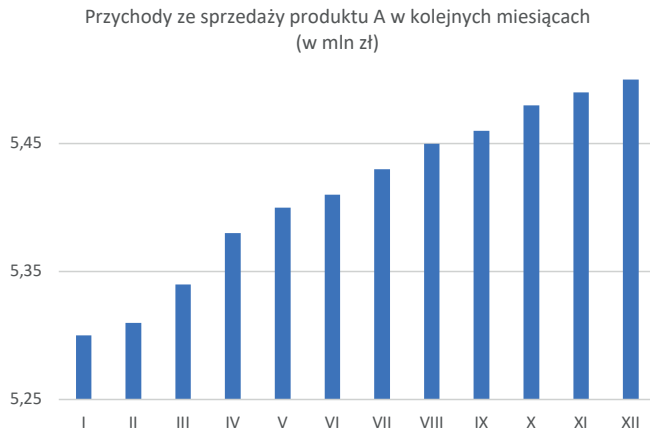


Wykres 3.2. Deformacja wykresu 3.1 – przykład 1

Źródło: opracowanie własne.

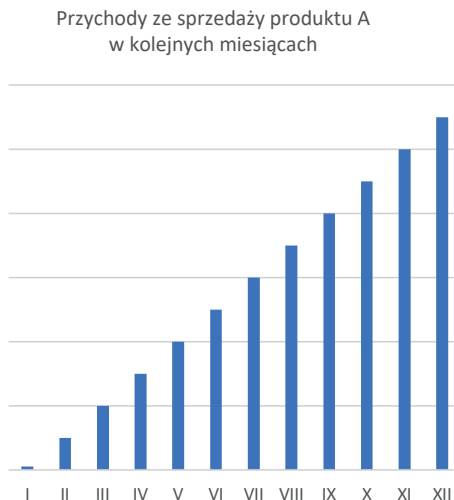
Równie korzystnie dla producenta produktu A dane będą prezentowały się na wykresie kolumnowym o skali osi Y rozpoczynającej się w punkcie minimalnie niższym niż najniższa wartość przychodów ze sprzedaży i osiagającej swoje maksimum w punkcie minimalnie niższym od najwyższej wartości przychodów (wykres 3.3).

W ostateczności, gdy wzrost przychodów jest coraz wolniejszy, można zastosować jeszcze bardziej nieetyczne rozwiązanie i w ogóle zrezygnować ze skali na osi Y, a wykres dodatkowo rozciągnąć w pionie (wykres 3.4) – zamieni się on wtedy w infografikę, której jedynym celem jest przekaz – „przychody ze sprzedaży cały czas rosną”. Owszem, przedstawiony trend będzie zgodny z rzeczywistością, jednak jego siła już nie.



Wykres 3.3. Deformacja wykresu 3.1 – przykład 2

Źródło: opracowanie własne.



Wykres 3.4. Deformacja wykresu 3.1 – przykład 3

Źródło: opracowanie własne.

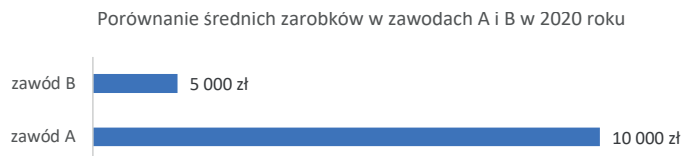
Bardzo często wykorzystywanym manewrem przez projektantów wykresów jest ustalanie początku osi Y nie w punkcie 0, ale na arbitralnie wybranej wartości. Porównując paski na wykresie (kolumnowym lub słupkowym), odbiorca intuicyjnie porównuje proporcje długości. Jeśli jeden pasek jest dwukrotnie dłuższy od drugiego, to z założenia odpowiada dwukrotnie większej wartości. Gdy paski nie są przedstawiane w całości, a punkt 0 znajduje się poza wykresem, to pole do manipulacji jest bardzo duże. Przesuwanie początku paska w kierunku wartości dodatnich optycznie zwiększa różnicę pomiędzy wielkościami, a przesuwanie w kierunku wartości ujemnych tę różnicę wizualnie zmniejsza.

Na wykresie 3.5 oba słupki zaczynają się (tak jak być powinno) w punkcie 0, na wykresie 3.6 w punkcie 4000, a na wykresie 3.7 w punkcie -4000. Za każdym razem odbiorca ma inne wrażenie dotyczące dwóch porównywanych wielkości, jakimi są tu średnie zarobki w zawodach A i B.



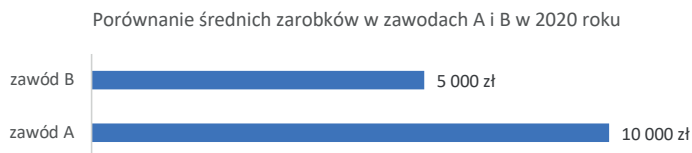
Wykres 3.5. Porównanie średnich zarobków w zawodach A i B

Źródło: opracowanie własne.



Wykres 3.6. Deformacja wykresu 3.5 – przykład 1

Źródło: opracowanie własne.



Wykres 3.7. Deformacja wykresu 3.5 – przykład 2

Źródło: opracowanie własne.

Lie factor, czyli „współczynnik kłamstwa” zdefiniowany przez Edwarda Tuftego (2007) pozwala ocenić jak bardzo wykres przekłamuje rzeczywiste relacje w danych. Opisuje on zależność między wielkością efektu pokazanego na grafice a wielkością efektu wynikającego z danych, na których bazuje grafika. Reprezentacja liczb, mierzona fizycznie na powierzchni grafiki, powinna być wprost proporcjonalna do rzeczywistych wartości.

$$Lie\ factor = \frac{\text{wielkość efektu pokazanego na grafice}}{\text{wielkość efektu wynikająca z danych}} \quad (3.1),$$

gdzie:

$$\text{wielkość efektu} = \frac{\text{druga wartość} - \text{pierwsza wartość}}{\text{pierwsza wartość}} \quad (3.2),$$

przy czym wielkość efektu traktowana jest tu jako zmiana procentowa.

Aby grafika nie zniekształcała odbioru, *lie factor* powinien mieć wartość między 0,95 a 1,05. Tak drobne różnice w wizualnym przekazie, zdaniem twórcy współczynnika, nie zakłóca odbioru. Jeśli jednak wartość *lie factor* jest mniejsza lub większa, oznacza to pewną (często zamierzoną) manipulację, mającą na celu zaniżenie lub zawyżenie rzeczywistych danych.

Warto dodać, że algorytm wyznaczania wielkości efektu sformułowany przez Tuftego wywołuje dyskusje naukowców, którzy sposób liczenia zmiany procentowej (polegający na dzieleniu zmiany przez wartość początkową) uważają za wadliwy i proponują własne formuły (https://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=00002o). Mimo różnych wzorów cel jest ten sam – wykazanie rozbieżności między zmianami wielkości na grafice a zmianami realnych danych.

Zgodnie ze wzorem 3.1, dla wykresu 3.6 *lie factor* wynosi:

$$Lie\ factor = \frac{\frac{6-1}{1}}{\frac{10000-5000}{5000}} = \frac{5}{1} = 5$$

a dla wykresu 3.7:

$$Lie\ factor = \frac{\frac{14-9}{9}}{\frac{10000-5000}{5000}} = \frac{0,56}{1} = 0,56,$$

czyli wykres 3.6 mocno zawyża i uwypukla różnice, a wykres 3.7 zaniża dysproporcje między faktycznymi zarobkami w zawodach A i B.

Jako manipulację skalą potraktować należy również niezachowywanie proporcji podczas prezentowania danych w postaci pól. Odbiorcom jeszcze trudniej zauważyć zniekształcenie niż w przypadku wykresów słupkowych czy kolumnowych. Oko ludzkie słabo radzi sobie z porównywaniem długości, ale jeszcze gorzej z analizą wielkości pól. Im bardziej nieregularny kształt, tym trudniej ocenić jego pole. Utrudnia to wykrywanie zniekształceń dotyczących przedstawiania pól w innych proporcjach niż te wynikające z danych. Ocena wielkości powierzchni podatna jest ponadto na sugestie związane z jasnymi intensywnymi kolorami, które optycznie powiększają pola (Biecek, 2014).

Założmy, że pewna organizacja chce zachęcić do udziału w konkursie i prezentuje dane z poprzedniej jego edycji. Zasady są takie, że nagrody o wartości powyżej 100 zł wygrywa 2% startujących. Poprawnie zaprezentowano to na wykresie 3.8. Projektant przygotował jednak grafikę (wykres 3.9), na której stosunek wygrywających wartościowe nagrody do wszystkich biorących udział w konkursie został zawyżony ponad czterokrotnie. Odbiorcom bardzo trudno dostrzec, że proporcje nie odpowiadają faktycznym danym.

Mała powierzchnia pomarańczowego pola nie zachęciłaby ludzi do udziału w konkursie tak mocno jak wizualizacja za pomocą zniekształconego wykresu.



Wykres 3.8. Zasady podziału nagród w konkursie

Źródło: opracowanie własne.

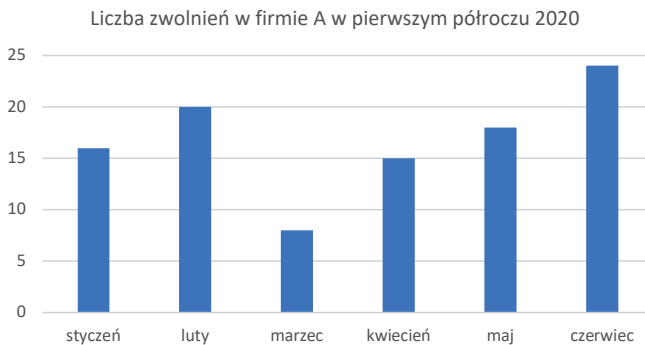


Wykres 3.9. Deformacja wykresu 3.8

Źródło: opracowanie własne.

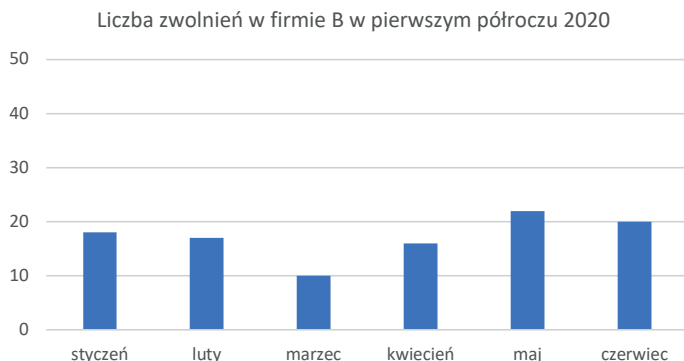
Innym przykładem manipulacji za pomocą zmiany skali jest zestawienie porównawcze tego samego typu danych, dotyczące np. dwóch regionów lub dwóch firm. Gdy projektant prezentuje dane na wykresach umieszczonych obok siebie, to zmieniając skalę na jednym z nich, może mocno zniekształcić przekazywane informacje i w znacznie lepszym świetle przedstawić jedną z firm na tle drugiej.

Pierwszy rzut oka odbiorcy na wykresy 3.10 i 3.11 powoduje odbiór komunikatu – „firma B zwalnia mniej pracowników”.



Wykres 3.10. Liczba zwolnień w firmie A

Źródło: opracowanie własne.



Wykres 3.11. Liczba zwolnień w firmie B

Źródło: opracowanie własne.

Dopiero bardziej szczegółowa analiza pozwala się zorientować, że skala zwolnień w firmie A i B jest porównywalna, a gdyby projektant zamieścił nad kolumnami etykiety danych, to po zsumowaniu wartości wręcz okazałoby się, że firma B w badanym okresie zwolniła więcej osób niż firma A.

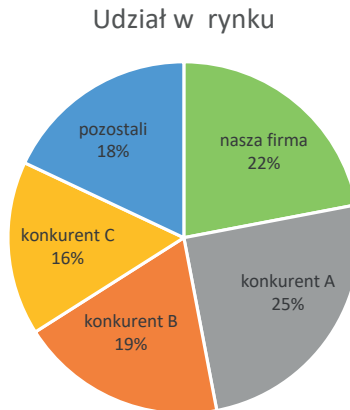
3.3. Perspektywa i kolorystyka

Kolejna iluzja optyczna pozwalająca na manipulację odbiorcą to zastosowanie efektu perspektywy w wykresach 3D. Najbardziej podatne na ten zabieg są wykresy kołowe. Oprócz elementu perspektywy, kluczowy wycinek koła można umiejscowić tak, by znajdował się z przodu wykresu. Ustawienie odpowiedniego kąta nachylenia i kolorystyki sprawi, że odbiór wizualizacji będzie taki, jak zaplanował jej projektant, mimo że dane wskazują na inny stan faktyczny.

Rzeczywistą sytuację firmy na tle konkurencji przedstawiono na wykresie 3.12. Na wykresie 3.13 firma wypada już zdecydowanie korzystniej, choć jej udział w rynku to 22%, podczas gdy konkurent A ma 25%. Pozytywny wizerunek firmy dodatkowo wzmocniono kolorem zielonym, najważniejszego konkurenta osłabiono kolorem szarym, a kolejnego konkurenta – negatywnie kojarzącą się barwą czerwoną.

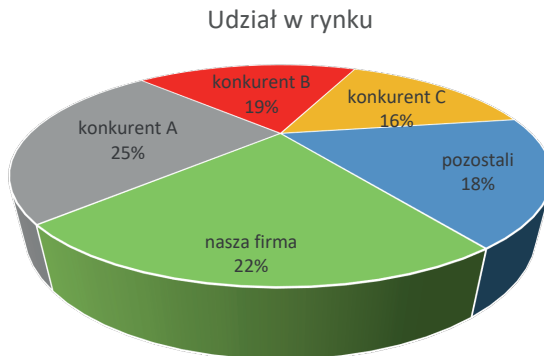
Trzeci wymiar i perspektywa całkowicie zakłóca odbiór wykresów kołowych, których ideą jest pokazanie wycinków koła jako fragmentów całości. Ważny jest zatem kąt – im większa jego wartość, tym większe pole wycinka. Zastosowanie perspektywy uniemożliwia wizualny odczyt proporcji, odbiorca ma do czynienia

z elipsą a nie kołem i tym samym zatarciu ulega rzeczywisty stosunek między fragmentami koła. Aby całkowicie pozbawić odbiorcę szans na zapoznanie się z faktyczną sytuacją, można jeszcze usunąć z wykresów etykiety danych w postaci procentowej.



Wykres 3.12. Udział w rynku

Źródło: opracowanie własne.



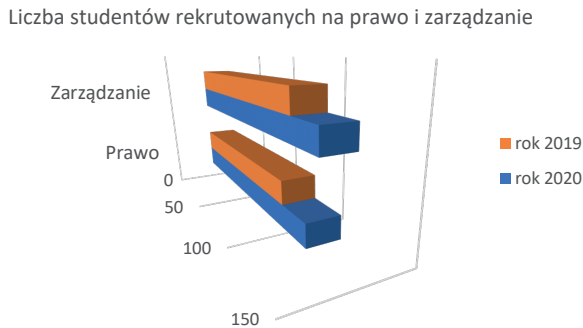
Wykres 3.13. Deformacja wykresu 3.12

Źródło: opracowanie własne.

3.4. Obrót

Oprócz perspektywy, przy tworzeniu wykresów trójwymiarowych, projektant ma opcję wyboru obrotów ścian wykresu względem osi X lub Y. Efekt ten nie jest elementem klasycznej manipulacji, ale zdecydowanie zaburza postrzeganie, wizualizacja staje się nieczytelna, utrudnione jest porównanie wartości na wykresie. Szews (2017) tego typu zabiegi nazywa „jałową estetyzacją ze zbyt dużą liczbą zbędnych i szkodliwych efektów”.

Wykres 3.14 mógłby wykonać projektant, który został poproszony o podanie liczby kandydatów przyjętych na studia na dwa różne kierunki i bardzo nie chce ujawnić, który kierunek cieszy się większym zainteresowaniem studentów. Trudno inaczej wytłumaczyć tak słabą i nieczytelną formę wizualizacji.



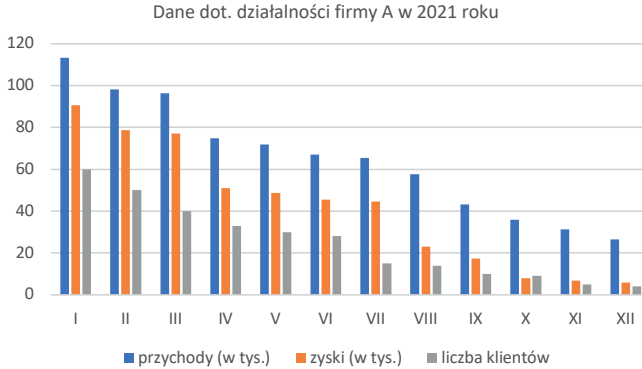
Wykres 3.14. Liczba rekrutowanych studentów

Źródło: opracowanie własne.

Połączenie efektu 3D z obrotem, przy jednoczesnym braku etykiet danych, całkowicie utrudnia odczyt informacji dotyczących liczby kandydatów przyjętych na prawo i zarządzanie w kolejnych latach.

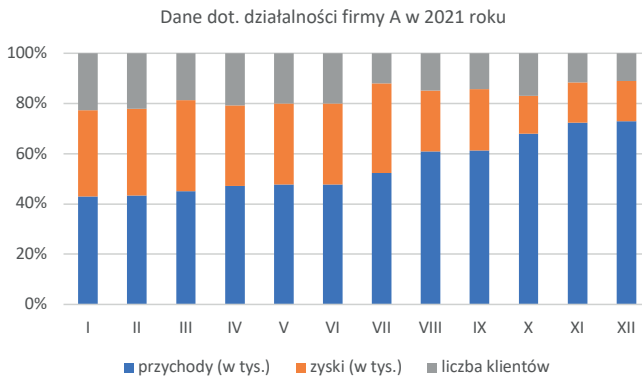
3.5. Ukrywanie trendów

Celowe ukrycie niewygodnych danych, na przykład kilku spadających, powiązanych ze sobą wskaźników, typu: przychody firmy, jej zyski i liczba klientów (wykres 3.15), osiągnąć można prezentując je na wykresie skumulowanym do 100% (wykres 3.16).



Wykres 3.15. Dane dotyczące działalności firmy A

Źródło: opracowanie własne.



Wykres 3.16. Dane z wykresu 3.15 skumulowane do 100%

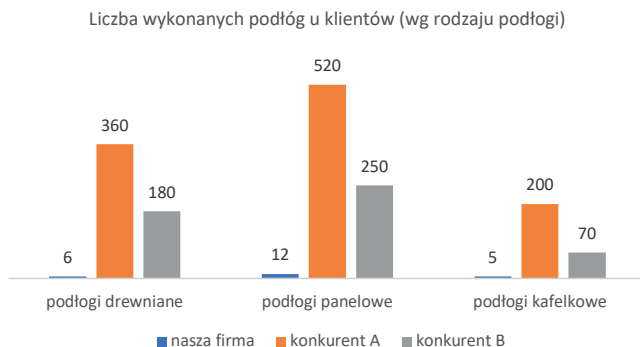
Źródło: opracowanie własne.

Taka wizualizacja znacznie utrudnia odbiorcy dostrzeżenie negatywnej sytuacji w firmie. Analiza udziału procentowego poszczególnych kategorii w całości wymaga większego skupienia i głębszego wnioskowania. Na wykresie 3.16 przychody prezentują się rosnąco, choć tak naprawdę to rośnie ich udział procentowy względem pozostałych dwóch wielkości. Skoro odsetek przychodów rośnie, a odsetek zysków spada, to znaczy, że rosną koszty. Malejąca liczba klientów również wskazuje na kłopoty firmy, choć w teorii (gdyby mniej klientów zapewniało taki sam przychód jak ich większa liczba) nie jest jednoznacznym wskaźnikiem złej sytuacji.

Intencją projektanta wykresu skumulowanego mogło być zatem celowe zatajenie informacji i sprawienie, by odbiorca zupełnie inaczej postrzegał firmę.

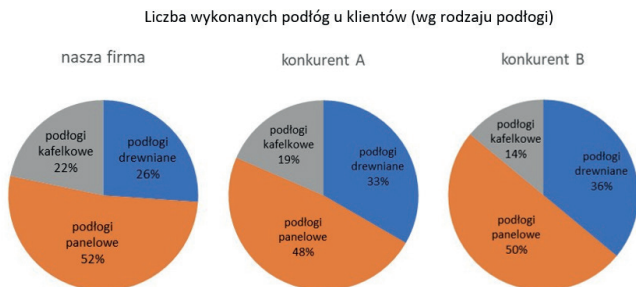
3.6. Dane procentowe zamiast danych bezwzględnych

Założmy, że wchodząca dopiero na rynek firma, ma niewielu klientów i wykonała niewiele zleceń. Chcąc jednak zachęcić do korzystania ze swoich usług, można zaprezentować się obok najbardziej znanych w branży, ale w zestawieniu pokazującym dane procentowe. Rzeczywiste dane dotyczące liczby zrealizowanych robót (wykres 3.17) nie przekonają tak szybko potencjalnego klienta do współpracy, jak zaprezentowanie tych samych danych jako odsetka wszystkich wykonanych zleceń (wykres 3.18).



Wykres 3.17. Liczba wykonanych podłóg w wartościach bezwzględnych

Źródło: opracowanie własne.



Wykres 3.18. Liczba wykonanych podłóg w ujęciu procentowym

Źródło: opracowanie własne.

Wykres 3.18 sugeruje wręcz, że „nasza” firma wypada lepiej od konkurencji, jeśli chodzi o najczęściej zamawiane podłogi panelowe. Ukrywając bezwzględna liczbę wykonanych zleceń łatwo wprowadzić w błąd klienta. Zestawienie procentowe wykonania poszczególnych rodzajów podłóg jedynie w ramach danej firmy nie daje bowiem możliwości porównania z konkurencją.

3.7. Punkt odniesienia

Kolejną formą zniekształcania informacji jest prezentowanie nie tych zależności, o które chodzi odbiorcy, ale innych, jednak tak blisko związanych z analizowanym zagadnieniem, że odbiorca może się nie zorientować, iż został wprowadzony w błąd.

Założmy, że kierownik działu handlowego pewnej firmy dysponuje następującymi danymi (tab. 3.1), które ma zaprezentować dyrekcji.

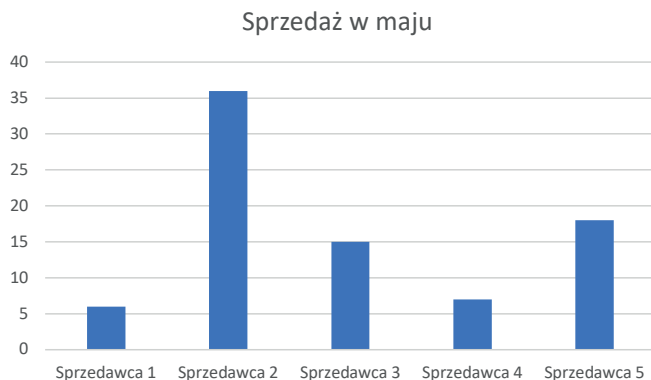
Tabela 3.1. Dane dotyczące sprzedaży

Pracownicy	Liczba sprzedanych sztuk	Cena jednostkowa (zł)	Wartość sprzedaży (zł)
Sprzedawca 1	6	170	1020
Sprzedawca 2	36	25	900
Sprzedawca 3	15	50	750
Sprzedawca 4	7	200	1400
Sprzedawca 5	18	60	1080

Źródło: opracowanie własne.

Gdy kierownik darzy sympatią sprzedawcę 2, ale widzi, że wartość sprzedanych przez niego towarów jest poniżej średniej, może przedstawić jego „zasługi”, ukazując na wykresie liczbę sprzedanych sztuk, a nie wartość sprzedaży (wykres 3.19).

Szybki rzut oka na wykres nie pozwoli dyrektorowi firmy zorientować się, że przedstawiono na nim inne dane, niż te, które są ważne z jego punktu widzenia. Dodatkowo utrudnia to braku opisu osi Y.



Wykres 3.19. Dane dotyczące sprzedaży

Źródło: opracowanie własne.

3.8. Prezentowanie wybiórczych danych

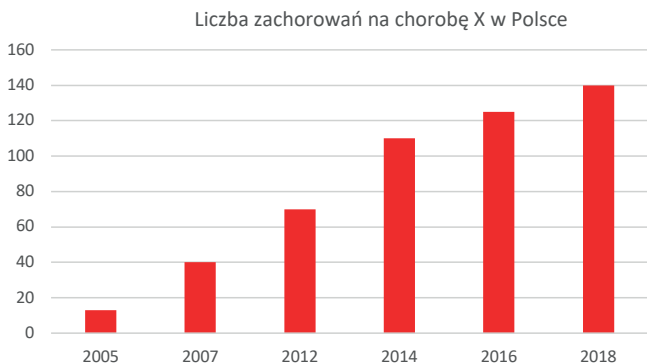
Ostatni z zaprezentowanych zabiegów manipulacyjnych jest ściśle powiązany z kolejnym wybiegiem, na jaki często pozwalają sobie projektanci grafik, chcąc skłonić odbiorcę do konkretnego toku rozumowania.

Na wykresie 3.20 przedstawiono liczbę zachorowań w Polsce na wirusową chorobę zakaźną X w kolejnych latach, poczynając od 2005 do 2018 roku. Zapobieganie chorobie polega na szczepieniach w wieku dziecięcym. W 2018 roku w stosunku do roku poprzedniego widoczny był wzrost liczby nowych przypadków. Chcąc mocniej zwrócić uwagę Polaków na problem i podkreślić znaczenie szczepień, jakim obowiązkowo należy poddawać dzieci, projektant wizualizacji postanowił wybrać i zaprezentować jedynie te dane, które będą dobrze pasowały do „opowiadanej historii” i celu wizualizacji (wykres 3.21). Przekaz ma być jednoznaczny – „szczepmy się, bo zapadalność na tę niebezpieczną chorobę szybko rośnie” i ma zrobić zamierzone wrażenie na odbiorcy.



Wykres 3.20. Zachorowania na chorobę X w Polsce

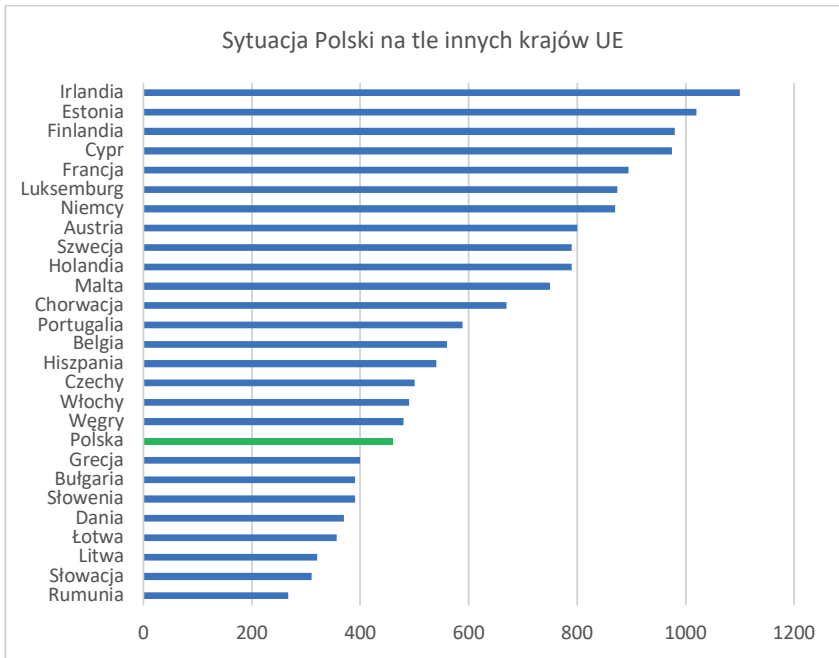
Źródło: opracowanie własne.



Wykres 3.21. Wybiórcze dane z wykresu 3.20

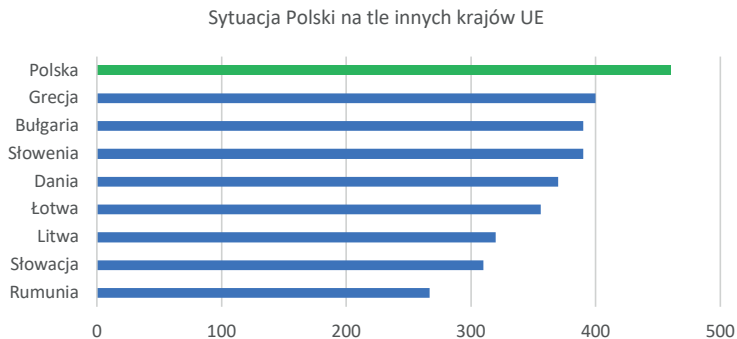
Źródło: opracowanie własne.

Poniżej omówiono kolejny przykład prezentowania selektywnych danych. Na wykresie 3.22 przedstawiono aktualną sytuację Polski w pewnym obszarze (im wyższe miejsce na liście tym lepiej). Wśród państw europejskich jesteśmy pod koniec rankingu. Projektant wizualizacji nie chcąc jednak tego ujawniać, może zaprojektować wizualizację (wykres 3.23), po obejrzeniu której odbiorca odniesie błędne wrażenie, jakoby nasz kraj przodował w zestawieniu. Wizualizacja nie kłamie, ale też, poprzez wybiórcze przedstawienie danych, zataja ważne fakty.



Wykres 3.22. Sytuacja Polski na tle krajów UE

Źródło: opracowanie własne.

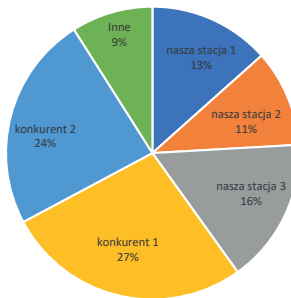


Wykres 3.23. Wybiórcze dane z wykresu 3.22

Źródło: opracowanie własne.

Inny sposób manipulacji, wynikający z wybiórczej prezentacji danych, może polegać na wykorzystaniu sytuacji i wybraniu takiego momentu, w którym konkurencja wypadła zdecydowanie gorzej. Mimo że zazwyczaj nas wyprzedza (wykres 3.24), to wyjątkowa sytuacja (np. strajk) może sprawić, że konkretnego dnia jej wskaźniki były dużo gorsze. Na przykład stacje telewizyjne, które nie mogą pogodzić się z niższą oglądalnością niż ma konkurencja, bardzo chętnie na wykonanie pomiarów oglądalności wykorzystają dzień, w którym konkurencja nie będzie nadawać. Okaże się, że wyniki będą znacznie lepsze niż normalnie, podczas cotygodniowych wcześniejszych pomiarów (wykres 3.25). Taką wizualizacją (bądź co bądź zgodną z prawdą, aczkolwiek nieodzwierciedlającą rzeczywistości) mogą się posługiwać w celach marketingowych.

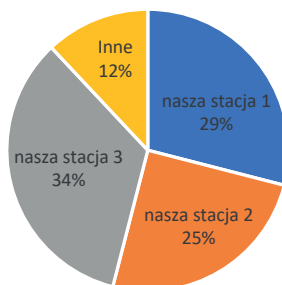
Średnia oglądalność w ostatnim roku



Wykres 3.24. Średnia oglądalność w ostatnim roku

Źródło: opracowanie własne.

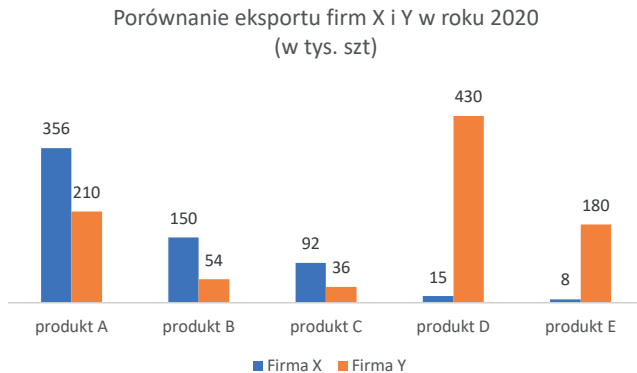
Oglądalność konkretnego dnia



Wykres 3.25. Oglądalność konkretnego dnia

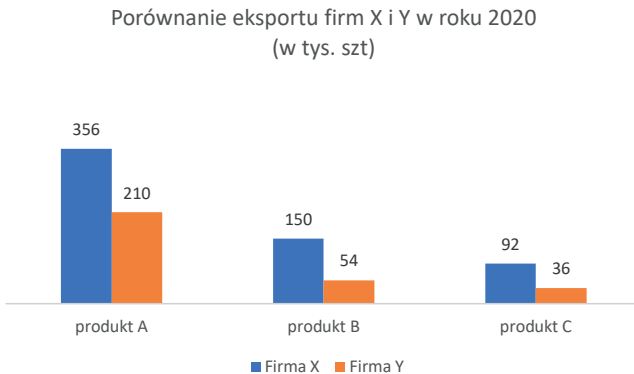
Źródło: opracowanie własne.

Poniżej jeszcze jeden przykład wybiórczej prezentacji danych, który całkowicie zniekształca faktyczny obraz rzeczywistości. Firma X i firma Y to jedni z największych eksporterów produktów A, B, C, D i E. Uczciwe zestawienie wyników eksportu obu firm zawierające wszystkie dane, tak, aby odbiorca mógł samodzielnie wyciągnąć wnioski zaprezentowano na wykresie 3.26. Gdy jednak projektant wizualizacji faworyzuje firmę X, to może przygotować wizualizację jak na wykresie 3.27 i wywoływać u odbiorcy błędne przekonanie o lepszej pozycji jako eksportera firmy X niż firmy Y.



Wykres 3.26. Porównanie eksportu firm X i Y

Źródło: opracowanie własne.

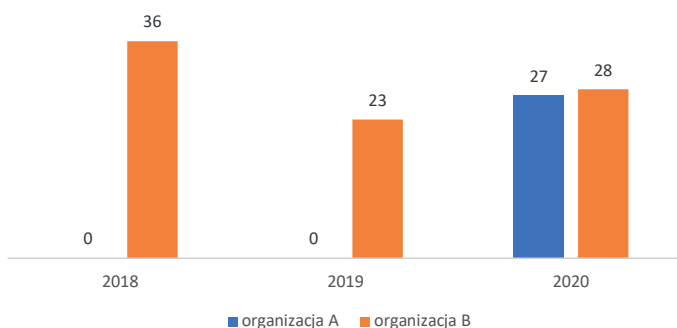


Wykres 3.27. Wybiórcze dane z wykresu 3.26

Źródło: opracowanie własne.

Ostatni przykład selektywnego traktowania danych, w tym przypadku również polegającego na zatajaniu niektórych faktów, zaprezentowano na wykresie 3.28. Wynika z niego, że organizacja A nie zrealizowała ani jednego projektu w 2018 i 2019 roku. Jej wyniki prezentują się zatem znacznie gorzej niż (faworyzowanej przez twórcę wykresu) organizacji B, która w latach 2018 i 2019 może pochwalić się licznymi aktywnościami, jeśli chodzi o działalność projektową.

Liczba projektów zrealizowanych przez organizacje A i B
w latach 2018–2020



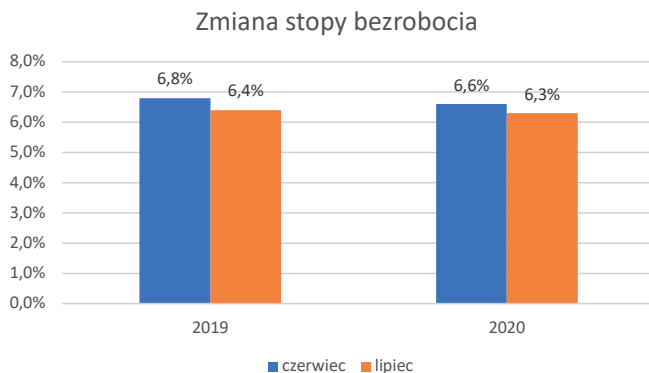
Wykres 3.28. Liczba projektów zrealizowanych przez organizacje A i B w latach 2018–2020

Źródło: opracowanie własne.

Fakty są jednak takie, że organizacja A powstała dopiero w roku 2020 i siłą rzeczy nie była w stanie wcześniej realizować żadnych projektów. Jest to na tyle ważna informacja, zwłaszcza gdy zestawienie dotyczy jedynie dwóch podmiotów, że w żadnym wypadku wizualizacja nie powinna obejmować lat wcześniejszych niż rok powstania organizacji A, nawet jeśli w tekście towarzyszącym wykresowi znajdzie się odpowiednie wyjaśnienie. Jest to świadome i celowe wprowadzanie w błąd odbiorców, którzy automatycznie gorzej ocenią młodą jednostkę, która dopiero zaistniała na rynku.

3.9. Niedozwolone operacje na osi czasu histogramu

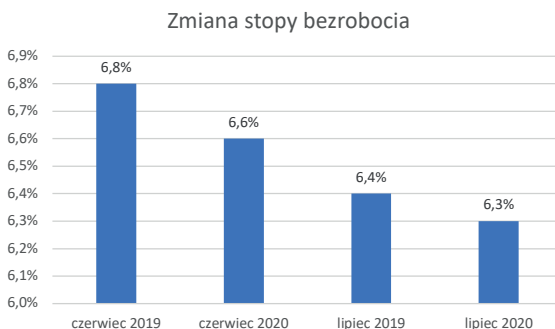
Poprawna wizualizacja informacji o zmianie stopy bezrobocia w dwóch miesiącach kolejnych lat powinna wyglądać tak jak na wykresie 3.29, z którego można odczytać, że utrzymuje się ona na prawie stałym poziomie.



Wykres 3.29. Zmiana stopy bezrobocia

Źródło: opracowanie własne.

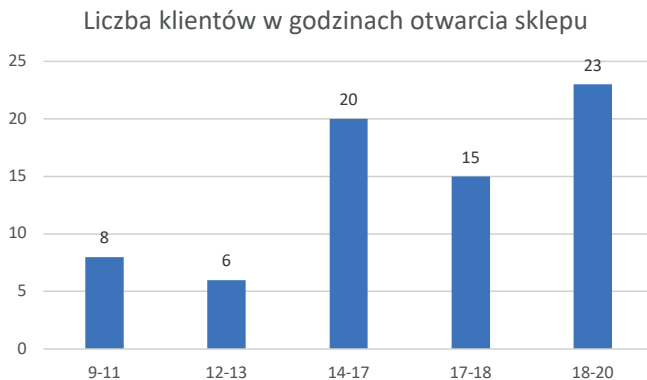
Projektantowi chodziło jednak o to, by w świadomości odbiorcy zakodowała się informacja, że stopa bezrobocia maleje. Celowo zatem tak rozmieścił kolejność pasków na wykresie (i dodatkowo nadał im jednakową kolorystykę oraz przyjął arbitralny początek osi Y), aby pokazać wyraźnie malejący trend (wykres 3.30). W rzeczywistości dokonał niedozwolonej operacji przestawiania pasków. Wykres ten nie jest bowiem wykresem słupkowym lecz histogramem, czyli wykresem, który służy do przedstawiania rozkładów zmiennych, a nie (jak wykresy słupkowe) do porównywania zmiennych. Histogramy obrazują dane ilościowe z zakresami danych pogrupowanymi w przedziały, a wykresy słupkowe prezentują dane kategoryjne (jakościowe). Paski wykresu słupkowego można dlatego dowolnie przestawiać (np. alfabetycznie, według długości słupka), natomiast układu pasków w histogramie nie można zmieniać, bo traci on sens.



Wykres 3.30. Deformacja wykresu 3.29

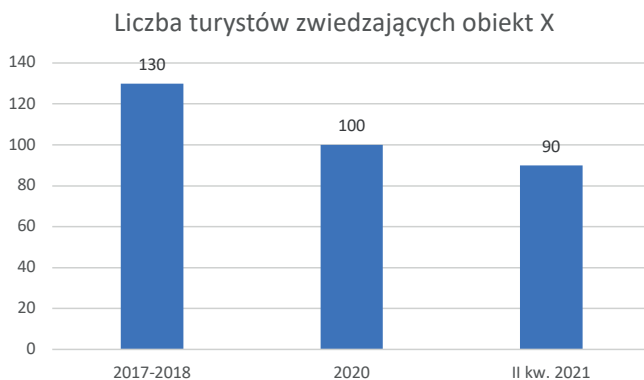
Źródło: opracowanie własne.

Innym przykładem (por. Cicharski, 2021) stosowania niedozwolonych operacji na histogramach jest podział na nierówne okresy (wykres 3.31) lub pomijanie niektórych okresów, innych łączenie w jeden, a jeszcze innych prezentowanie w sposób niepełny (wykres 3.32). Rozpiętość przedziałów w histogramie, w ramach których prezentowane są liczebności obserwacji, musi być równa.



Wykres 3.31. Liczba klientów w godzinach otwarcia sklepu – przykład deformacji histogramu

Źródło: opracowanie własne.



Wykres 3.32. Liczba turystów zwiedzających obiekt X – przykład deformacji histogramu

Źródło: opracowanie własne.

Celem niniejszego rozdziału była analiza różnych form manipulacji, jakich dopuszczają się projektanci wykresów. Omówiono zarówno zniekształcenia wynikające z iluzji optycznych, możliwych do osiągnięcia za pomocą wykorzystywanych programów komputerowych, służących do tworzenia wykresów, jak i wypaczenia powstałe przez celowy dobór danych, nie w pełni odzwierciedlających rzeczywistość. W praktyce bywają przypadki nieświadomych deformacji, będących skutkiem braku wiedzy i umiejętności sporządzenia odpowiedniej wizualizacji lub po prostu nieprzemyślenia skutków zastosowania takiego lub innego efektu graficznego. Sarah Leo (2019), przygotowująca wizualizacje dla *The Economist*, dokonała przeglądu różnego rodzaju błędów, które zdarzyło się popełnić redakcji, ale jak zapewnia, absolutnie nie było to działanie celowe.

Zdarzają się jednak przypadki umyślnego dezinformowania odbiorcy, do których dochodzi zarówno w sferze politycznej, jak i gospodarczej. Mogą być one efektem np. pewnej indoktrynacji prowadzonej przez rządzących, potrzeby pozyskania wyborców, za wszelką cenę przekonania do swoich racji. W biznesie zaś mogą wynikać np. z walki przedsiębiorstw o rynek, z chęci pokonania konkurencji i pozyskania większej liczby klientów. Obszarem, w którym tego typu świadome i zamierzone manipulacje zdarzają się dość często jest reklama i marketing. Projektantowi takich wizualizacji nie zależy na obiektywnym przekazaniu informacji, lecz na „opowiedzeniu pewnej historii”, która ukierunkuje sposób myślenia odbiorców i podsunie im pożądane wnioski.

Bardzo więc ważne jest zarówno kształcenie w zakresie rzetelności przekazu informacji oraz etyki dziennikarskiej, jak i podnoszenie poziomu świadomości w społeczeństwie w obszarze związanym z odbiorem informacji oraz wzmożona czujność przy interpretacji spotykanych w mediach wizualizacji. Zwłaszcza dziś, w dobie ery internetu i *fake newsów*, gdy nadmiar informacji, jakimi ludzie są bombardowani w każdej minucie, nie pozwala na ich szybką weryfikację. Czytelnicy często jedynie prześlizgują się wzrokiem po nagłówkach i ilustracjach, a rutyna, światopogląd i intuicja dopowiadają im resztę. Nie jest to groźne, dopóki dotyczy błahych informacji. W tej samej jednak rzeczywistości musi umieć odnaleźć się również kadra zarządzająca, na co dzień otrzymująca na swoje biurka liczne raporty, wykresy, prezentacje dotyczące sytuacji gospodarczej, wyników finansowych, poczynąń konkurencji. Często, nie mając czasu dokładnie się z nimi zapoznać, prezesi firm jedynie pobieżnie je przeglądają, zwracając uwagę na rzucające się w oczy i zapadające w pamięć grafiki, bo przecież jeden obraz zastępuje tysiąc słów. W ten sposób wyrabiają sobie zdanie na dany temat, a następnie podejmują ważne decyzje biznesowe.

Autorka liczy na to, że przedstawione przykłady będą nie tylko przyczynkiem teoretycznym do rozwoju wiedzy w obszarze wizualizacji, ale również odegrają rolę praktycznego przewodnika i wyczulą kierowników i dyrektorów przedsiębiorstw na czyhające na nich manipulacje informacją.

Intencją autorki jest również zwrócenie uwagi na kształcenie w zakresie umiejętności tworzenia poprawnych wizualizacji. Temu właśnie zagadnieniu szerzej poświęcono kolejny rozdział książki.

4. Badanie umiejętności studentów w zakresie wizualnego przekazu informacji

4.1. Cel badania

W erze społeczeństwa cyfrowego wyraźnie odczuwalne są problemy związane z szybkim przyswajaniem nadmiaru informacji, które, w celu łatwiejszego odbioru, coraz częściej prezentowane są w formie graficznej.

Kluczowe stają się zatem umiejętności związane ze swoistym kodowaniem i dekodowaniem informacji, czyli zamianą opisu słownego na postać wizualną i odwrotnie – prawidłowym odczytaniem obrazów i interpretowaniem zawartych na nich treści. Dylak (2012) uznaje wręcz, że kompetencje wizualne (*visual literacy*) są niezbędne, aby właściwie funkcjonować we współczesnym świecie.

W obszarze zarządzania, na którym koncentruje się autorka publikacji, jest to szczególnie istotne dla przedsiębiorców, którzy jak wspomniano już w poprzednim rozdziale, ze względu na tempo pracy, często podejmują ważne decyzje biznesowe jedynie na podstawie skróconych raportów z licznymi wykresami i zapadającymi w pamięć grafikami. Widać zatem jak ważna jest rola osób, które przygotowują owe raporty i wizualizacje. Jak istotne jest to, by treść wyrażona za pomocą słów, była odpowiednio (w pełni a zarazem prosto i zrozumiale) przełożona na postać graficzną, której celem jest znaczna poprawa percepcji, przy jednocześnie minimalnej utracie wartości informacyjnej.

Wielu autorów (m.in. Munzner, 2009; Manovich, 2010; Chan, 2017; Taei, 2018; Falkowitz, 2019) podkreśla znaczenie roli projektanta, która, tak jak przedstawiono to w rozdziale pierwszym, zmienia się w zależności od tego, co ma zaprezentować. W przypadku wizualizacji danych, projektant pełni rolę analityka: bada, modyfikuje, rozważa i wypróbowuje różne scenariusze, aby zapewnić obiektywizm i pozwolić odbiorcom wizualizacji na samodzielne wyciąganie wniosków. Równie ważna jest jego rola w przypadku wizualizacji informacji (wyniku wstępnych

analiz danych), która tworzona jest dla określonego zbioru danych, w celu zwrócenia uwagi odbiorcy na te relacje, na których pokazaniu zależy projektantowi. Trudne zadanie stoi też przed projektantem wizualizacji wiedzy, której celem jest usprawnienie procesów komunikacyjnych i prezentacja wiedzy z wykorzystaniem dostępnych zasobów wizualnych w sposób bardziej zrozumiały i komunikowalny dla innych ludzi niż opis słowny.

Celem badań, których wyniki omówiono w niniejszym rozdziale, była analiza umiejętności przedstawienia w postaci wykresu informacji dostępnych jako opis słowny. W roli projektantów wizualizacji postawiono studentów kierunków z zakresu nauk społecznych, jako tych, którzy w przyszłości podejmą pracę w obszarze szeroko pojętego zarządzania, zarówno w sektorze prywatnym, jak i państwowym, w charakterze ekonomistów, menedżerów (lub ich asystentów), doradców biznesowych i innych specjalistów od gospodarki. Wcześniejsze rozważania wskazują, że w przypadku takich osób bardzo ważna dla efektywnej realizacji zadań zawodowych jest wiedza na temat wizualizacji oraz praktyczne umiejętności z tym związane. Badanie ukazało poziom przygotowania studentów do tworzenia wizualizacji. Analiza i ocena jakości stworzonych przez nich wykresów zasygnalizowała problemy, których obie strony komunikacji (nadawca i odbiorca) nie zawsze są świadome, a których rozwiązania wydają się konieczne do sprawnego i efektywnego funkcjonowania w coraz bardziej ugruntowującej się gospodarce wiedzy.

Wyniki tego badania autorka niniejszej publikacji zaprezentowała podczas 26. międzynarodowej konferencji „Knowledge-Based and Intelligent Information & Engineering Systems” (Krok, 2022).

4.2. Metodyka badań

Badanie wykonano w kwietniu i maju 2021 roku z wykorzystaniem platformy MS Teams, w trakcie zajęć ze studentami, prowadzonych w formie zdalnej. Na przeprowadzenie badań wybierano taki rodzaj zajęć, podczas których studenci pracowali przy komputerach.

Ankiety skierowano do 142 osób, zwrotnie otrzymano 99 odpowiedzi, od 49 kobiet i 50 mężczyzn, z sześciu następujących kierunków studiów:

- bezpieczeństwo wewnętrzne (studia I stopnia),
- nauki o polityce (studia I stopnia),
- ekonomiczno-prawny (studia I stopnia),
- gospodarka nieruchomościami (studia I stopnia),

- informatyka w biznesie (studia I stopnia),
- management instytucji publicznych i public relations (studia II stopnia).

Studentów szczegółowo poinformowano o celu badania. Mieli też możliwość zadania pytań w przypadku wątpliwości i potrzeby dodatkowych wyjaśnień. Czas na wypełnienie ankiety był w zasadzie nieograniczony. Studenci mogli to zrobić od razu w trakcie zajęć lub jeśli nie mieli takiej chęci lub nie dysponowali w danym momencie niezbędnym, ich zdaniem, sprzętem lub oprogramowaniem, bądź też chcieli się dłużej zastanowić i poświęcić więcej czasu na realizację zadania, mogli przesłać ankietę w ciągu kilku dni.

Główne zadanie polegało na zapoznaniu się z informacją słowną, wyrażoną w postaci jednego zdania, a następnie przedstawieniu jej w postaci wykresu. Informacja brzmiała następująco: „Firma X w kolejnych miesiącach pierwszego półrocza 2019 roku zatrudniła następującą liczbę nowych pracowników: 10, 9, 12, 18, 20, 35, a w kolejnych miesiącach pierwszego półrocza roku 2020 odpowiednio: 8, 13, 12, 25, 30, 36”.

O typie, układzie, kolorystyce i elementach formatowania wykresu studenci mieli zdecydować samodzielnie. W zależności od własnych preferencji, wykres mogli przygotować w dowolnym programie komputerowym bądź ręcznie na kartce papieru (i załączyć plik w formacie właściwym dla użytego programu, zrzut ekranu, skan lub zdjęcie).

Kolejne pytania ankiety dotyczyły wykonanego wykresu. Studenci mieli podzielić się informacją, czy ich wykres wygląda tak, jak tego chcieli, a jeśli nie, to dlaczego. Mieli również podać, ile czasu zajęło im sporządzenie wykresu oraz subiektywnie ocenić, czy zadanie sprawiło im trudność. Ostatnie dwa pytania dotyczyły oceny przez studentów swoich umiejętności tworzenia wykresów oraz potrzeby poszerzenia wiedzy w tym zakresie.

4.3. Analiza przygotowanych przez studentów wizualizacji

Otrzymano pliki w różnych formatach: 43 w formacie xlsx, 31 w formacie png, 17 w formacie docx, 7 w formacie jpg i 1 plik pdf. Pliki w formatach png, jpg, a także pdf w większości zawierały wycinek ekranu z wykresem wykonanym w arkuszu kalkulacyjnym, a format docx osadzony obiekt arkusza. Jedynie kilka osób wykonało zadanie w innym programie niż arkusz kalkulacyjny. Przesłano też dwa zdjęcia odręcznie sporządzonych rysunków. Można zatem uznać, że studenci właściwie dobrali narzędzie informatyczne do zadania, jakie mieli wykonać.

Wszystkie prezentowane w tym rozdziale wykresy wykonane zostały przez uczestników badania i zamieszczone w postaci rysunków, bez jakichkolwiek modyfikacji. Typ wybranych przez studentów form wizualizacji przedstawiono w tabeli 4.1.

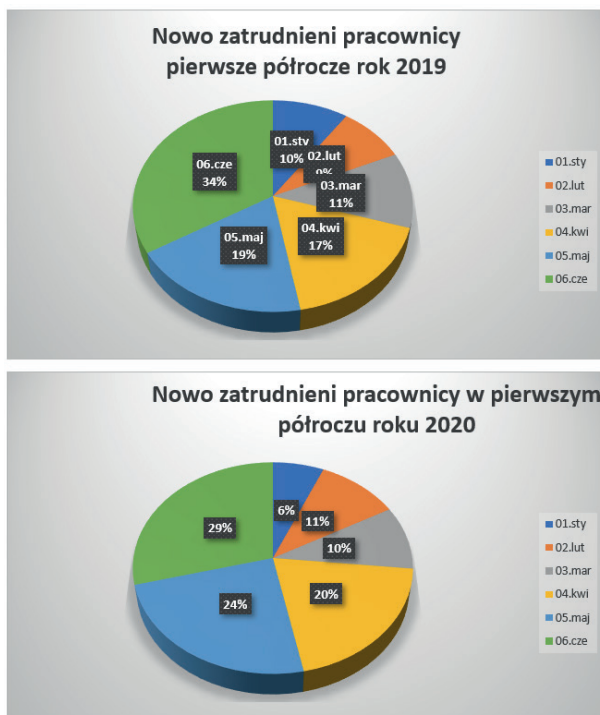
Tabela 4.1. Formy wizualizacji wybrane przez studentów

Typ wizualizacji	Liczba studentów
wykres kolumnowy	75
wykres słupkowy	6
wykres liniowy	7
tabela	9
wykres kołowy	1
wykres lejkowy	1

Źródło: opracowanie własne.

75 osób zdecydowało się na wykres typu kolumnowego. Dominowała forma najprostsza, czyli grupowana dwuwymiarowa (60), choć były też wykresy 3D (7) i skumulowane (2). 6 osób przedstawiło informację przy pomocy dwóch wykresów kolumnowych. Za bardzo zbliżoną można również uznać postać wykresu słupkowego, którą przygotowało 6 osób (w tym jeden wykres skumulowany i jedna wizualizacja z wykorzystaniem dwóch wykresów słupkowych). 10 osób wybrało formę tabelaryczną, mimo że w poleceniu wyraźnie była prośba o przedstawienie informacji w postaci wykresu. Stosunkowo niewiele, bo tylko 7 osób zaproponowało wykres liniowy, który pozwala na szybkie rozpoznanie tendencji danej zmiennej względem czasu (w tym przypadku liczby nowych zatrudnień w firmie X w pierwszych sześciu miesiącach dwóch analizowanych lat).

Jedna osoba zaprezentowała dane w postaci dwóch wykresów kołowych w ujęciu procentowym (rys. 4.1) (osobno dla 2019 r. i osobno dla 2020 r.), co, abstrahując od błędnej legendy, należy uznać za mało przejrzyste i nieadekwatne z punktu widzenia treści, jaką należało zwizualizować. Jedna osoba także zaprezentowała dane w postaci wykresu lejkowego (rys. 4.2), który może znaleźć zastosowanie w celu prezentacji wartości w przepływach i procesach, natomiast jest to nieprawidłowy wybór dla graficznego odwzorowania informacji w analizowanym przypadku.



Rysunek 4.1. Mało przejrzyste i nieadekwatne do prezentowanej treści wykresy kołowe z błędną legendą

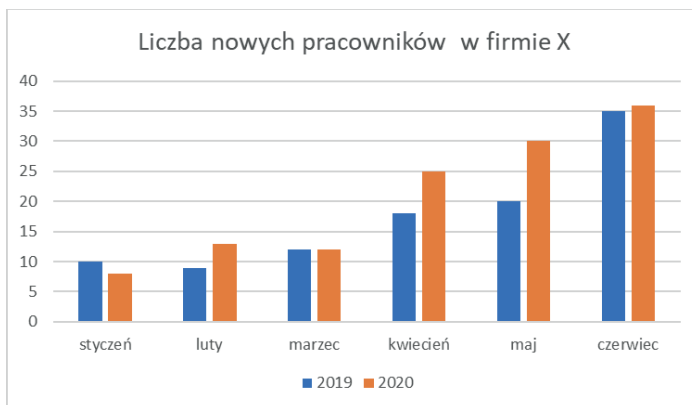
Źródło: uczestnik badania.



Rysunek 4.2. Nieprawidłowo dobrany do prezentowanych treści wykres lejkowy

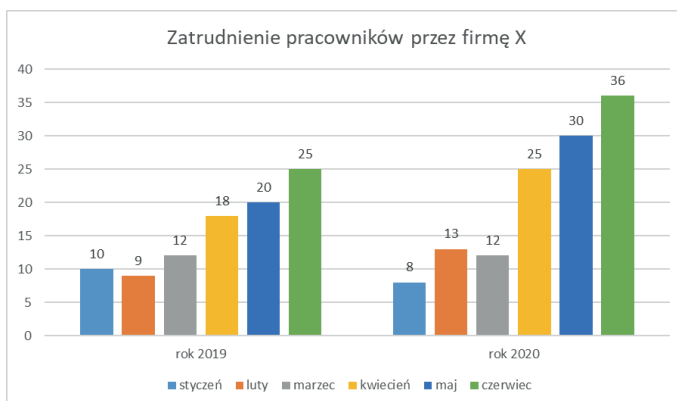
Źródło: uczestnik badania.

Jeśli chodzi o wykresy kolumnowe i słupkowe to dominował (42 wykresy) typ, w którym zestawiano obok siebie liczby zatrudnionych w poszczególnych miesiącach dla obu analizowanych lat (dwie serie danych dla poszczególnych miesięcy) (np. rys. 4.3). Drugim najpopularniejszym wariantem prezentacji (22 wykresy) było pokazanie danych o liczbie zatrudnionych w kolejnych sześciu miesiącach jednego roku, a następnie drugiego (sześć serii danych dla dwóch lat) (np. rys. 4.4), lub jako jedna seria danych (np. rys. 4.5).



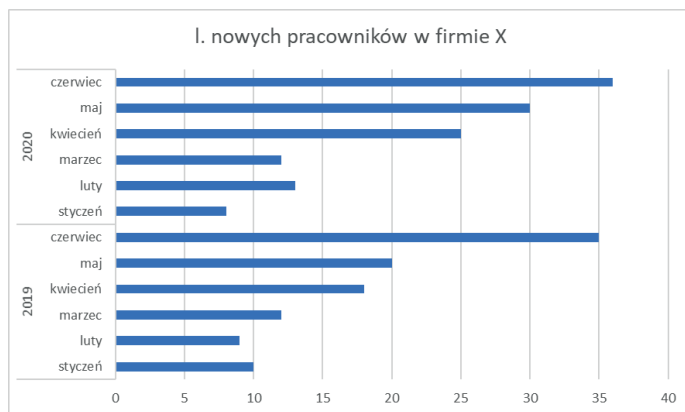
Rysunek 4.3. Wykres kolumnowy – dwie serie danych dla poszczególnych miesięcy

Źródło: uczestnik badania.



Rysunek 4.4. Wykres kolumnowy – sześć serii danych dla dwóch lat

Źródło: uczestnik badania.



Rysunek 4.5. Wykres słupkowy – jedna seria danych

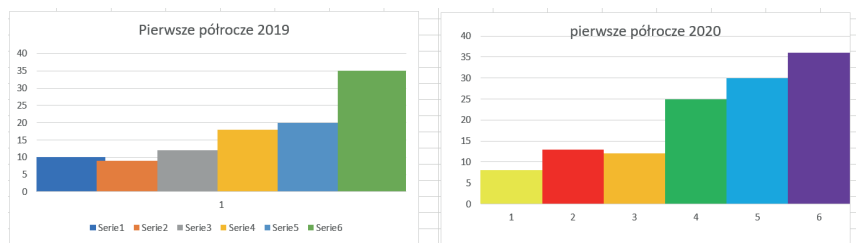
Źródło: uczestnik badania.

W większości sporządzonych wizualizacji nie modyfikowano domyślnego układu kolorystycznego zaproponowanego przez arkusz kalkulacyjny. 23 osoby zdecydowały się na wprowadzenie dodatkowych barw, wypełniając obszar wykresu, obszar kreślenia, nadając kolor liniom siatki, zmieniając wypełnienie pełne serii danych na gradientowe lub też formatując kolor czcionki wybranych tekstów.

Kolejny element, jaki poddano analizie to tytuł wizualizacji. 75 osób zatytułowało stworzoną przez siebie wizualizację, przy czym w 11 przypadkach tytuł był niepełny, to znaczy nie pozwalał (nawet przy uwzględnieniu pozostałych elementów na wykresie, takich jak wartości na osiach, ich opisy a także legenda) na odczytanie kompletnej informacji, jaka była zawarta w postaci słownej i jaka miała być zwizualizowana. Za niepełne uznano tytuły takie jak np.: „Pracownicy”, „Firma X”, „Zatrudnienia”, „Pierwsze półrocze 2019”. W czterech przypadkach tytuły były niepoprawne w kontekście wykresu (wykres nie przedstawiał tego, co sugerował tytuł lub tytuł nie miał związku z wykresem). Warto też wspomnieć, że w tytułach (mimo że przeważnie były to krótkie równoważniki zdań) pojawiło się wiele literówek oraz potknięć stylistycznych (np. „Liczba zatrudnień w danym okresie czasu”). Nagminnym błędem było też użycie słowa „ilość” zamiast „liczba” w odniesieniu do rzeczowników policzalnych (jakimi są nowo zatrudnieni). 17 osób nie nadało żadnego tytułu przygotowanej przez siebie wizualizacji, a 7 – pozostawiło domyślny napis, jaki pojawia się podczas tworzenia wykresu („tytuł wykresu”, „wykres”, „kolumna 2”).

24 osoby umieściły nad kolumnami lub wewnątrz kolumn etykiety danych, co znacznie ułatwia odczyt dokładnych wartości; 7 z nich dodatkowo zdecydowało się również na opisy osi (X i Y lub tylko jednej z nich), co także przyspiesza analizę prezentowanej treści.

Jeśli chodzi o legendę, to ten element wykresu zastosowało 66 badanych, przy czym 4 zrobiło to w błędny lub nieczytelny sposób (np. rys. 4.6). 33 osoby nie zastosowały legendy (przy czym w przypadku 9 jest to usprawiedliwione tabelaryczną formą prezentacji).



Rysunek 4.6. Niepoprawnie sporządzone wykresy kolumnowe – niewnosząca żadnych treści legenda

Źródło: uczestnik badania.

Po analizie graficznej (wizualnej) poniżej dokonano analizy merytorycznej wykresów, czyli czy faktycznie odzwierciedlają one tę informację, na bazie której zostały przygotowane.

Wyjściowa informacja słowna, która miała zostać zwizualizowana mówi o:

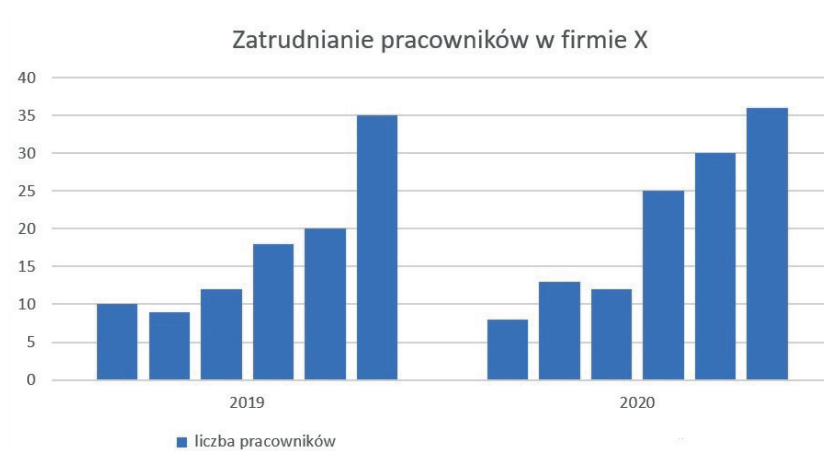
- firmie X,
- liczbie nowo zatrudnionych w niej pracowników,
- okresie zatrudniania, tj. w kolejnych miesiącach pierwszego półrocza 2019 i 2020 roku.

Gdy przygotowane przez studentów wizualizacje przeanalizuje się pod tym kątem, tj. czy zawierają wszystkie wyżej wymienione trzy elementy, to okazuje się, że jedynie 10 wykresów zawiera pełną informację, która miała być zwizualizowana.

To, że chodzi o firmę X zawarto na 49 wykresach.

Informacja, że chodzi o nowo zatrudnionych pracowników dostępna jest tylko na 25 wykresach. W większości wizualizacji (aż w 53) jest mowa po prostu o liczbie pracowników lub zatrudnionych, co w oczywisty sposób zniekształca przekaz.

Połączenie tych dwóch elementów, czyli, że chodzi o nowo zatrudnionych pracowników w firmie X widoczne jest jedynie na 14 wizualizacjach. Na dwóch z nich pominięto jednak informację o tym, że chodzi o zatrudnienie w kolejnych miesiącach pierwszego półrocza 2019 i 2020 roku (np. rys. 4.7), a na kolejnych dwóch błędnie zobrazowano okres zatrudniania.



Rysunek 4.7. Brak informacji o miesiącach, w jakich zatrudniano pracowników

Źródło: uczestnik badania.

W uzupełnieniu warto dodać, że poprawna informacja o okresie zatrudniania (kolejne miesiące pierwszego półrocza 2019 i 2020 r.) widoczna jest jedynie na 66 wizualizacjach. 16 studentów błędnie zaprezentowało okresy – przeważnie od stycznia do grudnia jednego roku (np. rys. 4.8 lub 4.9) lub sześć miesięcy jednego roku i inne sześć drugiego (np. rys. 4.10), a 17 osób w ogóle pominęło te kluczowe informacje (np. wspomniany już rys. 4.6, 4.7 lub 4.11).



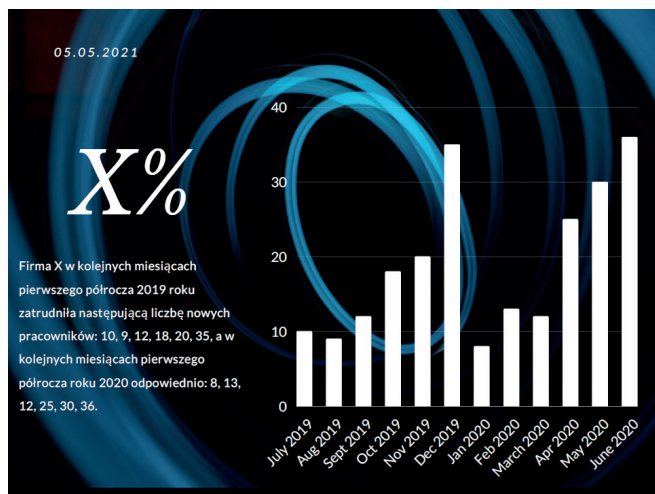
Rysunek 4.8. Błędnie przedstawiony okres zatrudniania – 12 miesięcy jednego roku

Źródło: uczestnik badania (zachowano oryginalną pisownię).



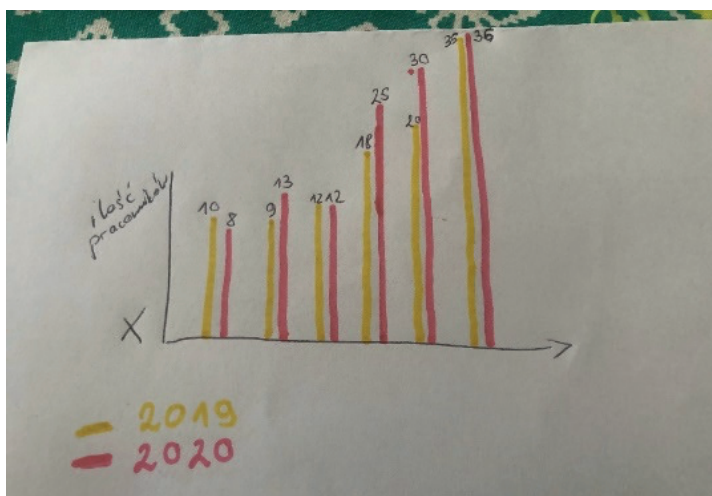
Rysunek 4.9. Błędnie przedstawiony okres zatrudniania, błędny tytuł i legenda wykresu

Źródło: uczestnik badania.



Rysunek 4.10. Błędnie przedstawiony okres zatrudniania – sześć ostatnich miesięcy 2019 roku i sześć pierwszych kolejnego roku

Źródło: uczestnik badania.



Rysunek 4.11. Brak nazw miesięcy, w których następowało zatrudnianie oraz brak informacji, że wykres przedstawia liczbę nowo zatrudnionych pracowników (a nie pracowników ogółem)

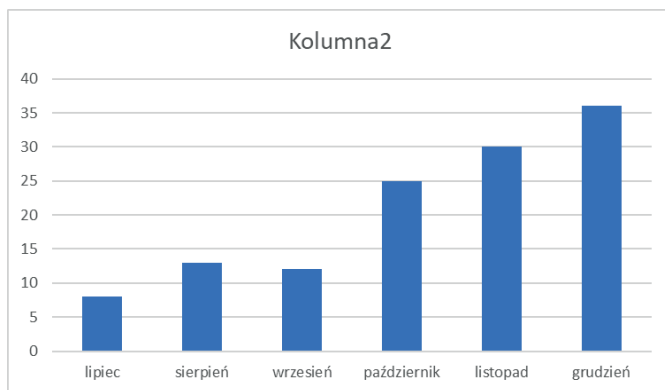
Źródło: uczestnik badania.

Tym samym można uznać, że jedynie 10% wykonanych wizualizacji jest poprawnych, tzn. w pełni zaprezentowano na nich tę informację, która miała być przekazana (np. rys. 4.3), wśród nich znalazło się: 6 wykresów kolumnowych, 1 słupkowy, 2 liniowe i 1 tabela.

Wszystkie wizualizacje, które prezentowały pełną informację miały tytuł i tylko jedna z nich pozbawiona była legendy. Niestety na żadnym z tych wykresów nie zamieszczono etykiet danych (wartości nad kolumnami), pozwalających na szybkie i precyzyjne odczytanie dokładnej liczby zatrudnionych w danym miesiącu, bez konieczności porównywania wysokości kolumny lub długości słupka ze skalą na osi. Wpływa to negatywnie na czytelność i przystępność wizualizacji, zwłaszcza w przypadku zainteresowania odbiorcy szczegółowymi danymi.

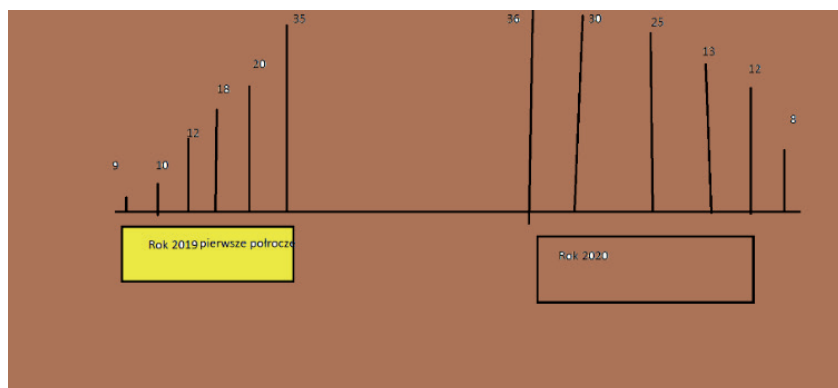
Podsumowując, najczęstsze błędy popełnione przez studentów, powodujące, że wizualizacja jest niepoprawna lub fragmentaryczna bądź zdecydowanie zmniejszające jej czytelność i utrudniające wnioskowanie to:

1. Prezentacja samych danych liczbowych, z pominięciem niezbędnych opisów (w postaci np. kategorii na osi X, legendy bądź tytułu), co uniemożliwia odczytanie właściwej informacji, bo odbiorca nie ma szansy odgadnąć czego dotyczą dane (np. rys. 4.12, 4.13, 4.14).
2. Prezentacja niewłaściwych danych – dotyczyło to głównie analizowanego okresu. Na wielu wykresach pojawiały się miesiące od stycznia do grudnia w ramach jednego roku (np. rys. 4.9).
3. Pominięcie niektórych danych (np. rys. 4.15).



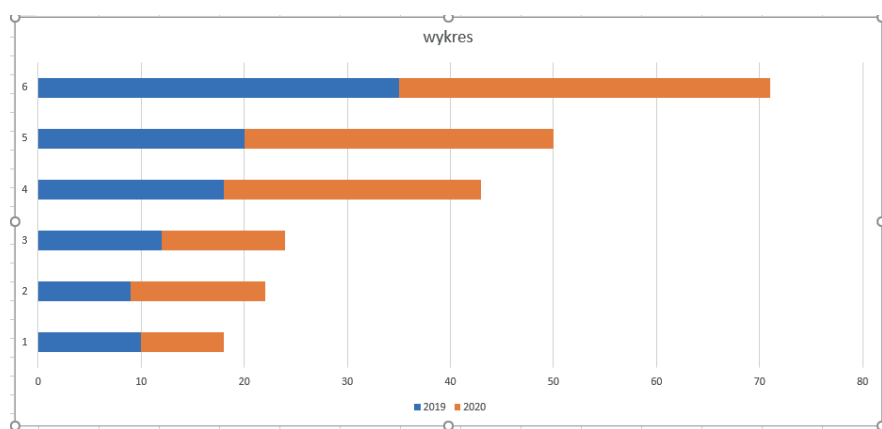
Rysunek 4.12. Brak informacji co zaprezentowano na wykresie – przykład 1

Źródło: uczestnik badania.



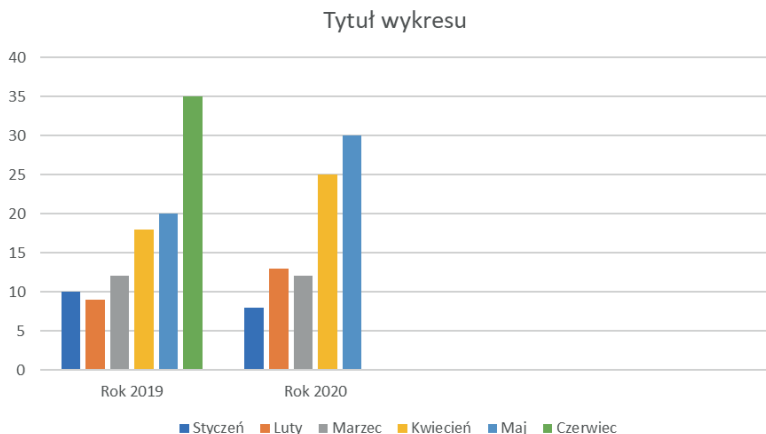
Rysunek 4.13. Brak informacji co zaprezentowano na wykresie – przykład 2

Źródło: uczestnik badania.



Rysunek 4.14. Brak informacji co zaprezentowano na wykresie – przykład 3

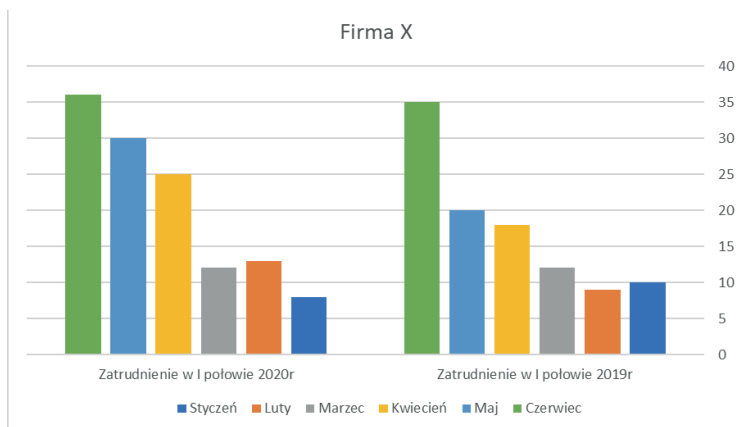
Źródło: uczestnik badania.



Rysunek 4.15. Pominięcie niektórych danych – brak danych dla czerwca 2020 roku

Źródło: uczestnik badania.

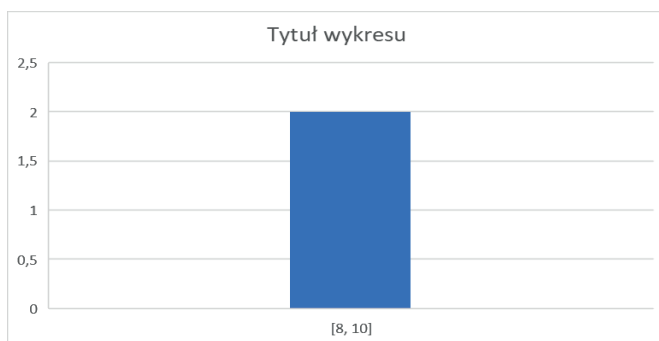
4. Brak etykiet danych (np. rys. 4.3, 4.5).
5. Prezentacja danych w układzie innym niż chronologiczny (np. najpierw dane za rok 2020 i to w odwrotnej kolejności, tj. od czerwca do stycznia, a później rok wcześniejszy w takim samym nieintuicyjnym układzie miesięcy) (np. rys. 4.16).



Rysunek 4.16. Niechronologiczny układ prezentacji danych utrudniający zrozumienie przekazywanych treści

Źródło: uczestnik badania.

6. Powielanie tych samych informacji w tytule wykresu i legendzie lub opisie osi (np. rys. 4.9).
7. Pozostawienie automatycznie wstawionych przez arkusz kalkulacyjny elementów (np. rys. 4.12, 4.14, 4.15, 4.17) zamiast ich usunięcia lub uzupełnienia właściwą treścią.
8. Umieszczenie pełnej informacji w postaci słownej obok wykresu prezentującego zupełnie inny przekaz (np. rys. 4.10).
9. Utworzenie wykresu w żaden sposób nienawiązującego do informacji, jaka była do przekazania (np. rys. 4.17).



Rysunek 4.17. Brak jakiegokolwiek związku wykresu z informacją, jaką miał prezentować

Źródło: uczestnik badania.

Można podejrzewać, że część z tych błędów wynika z braku wiedzy o tym, jak powinna wyglądać dobra wizualizacja i na jakie elementy należy zwrócić uwagę. Studenci mogli nie zadać sobie pytania, czy na podstawie wykresu bądź tabeli, które sporządzili, odbiorca będzie w stanie zbudować/odtworzyć pierwotne zdanie (przekazywaną informację). Niektóre z wykresów sugerują, że inną przyczyną błędów jest też brak dostatecznych umiejętności obsługi narzędzia informatycznego, w którym przygotowywane były wizualizacje. Problemy te przybliży analiza odpowiedzi na kolejne pytania ankiety.

4.4. Informacje uzupełniające badanie

Jak już wspomniano, po przygotowaniu wizualizacji studentów poproszono o przekazanie dodatkowych informacji, dotyczących między innymi zadowolenia z przygotowanego wykresu, opinii na temat stopnia trudności zadania, czasu jaki im ono zajęło oraz oceny własnych umiejętności w tym zakresie.

Na pytanie „czy Twój wykres wygląda tak, jak byś chciał(a)?” aż 80 osób odpowiedziało, że tak. Nie mieli żadnych uwag ani wątpliwości co do ostatecznej postaci wizualizacji. Deklarowany średni czas utworzenia wykresu w tej grupie to 6,8 min. Wśród tych osób większość, bo aż 62 studentów uznało, że przygotowanie wykresu nie stanowiło dla nich żadnego problemu, stopień trudności stworzenia wizualizacji ocenili na łatwy lub bardzo łatwy.

W grupie zadowolonych z opracowanej wizualizacji 10 osób swoje umiejętności tworzenia wizualizacji oceniła na duże/spore, a 47 osób na średnie. Co zwraca jednak uwagę, aż 22 osoby uznały, że mają niewielkie umiejętności w tym zakresie, a jedna osoba oceniła je na prawie zerowe. Pozostałe 19 osób, które przyznały, że stworzony przez nich wykres nie jest zgodny z ich oczekiwaniami, wskazały że:

- intuicyjnie czują, że wykres powinien wyglądać inaczej, ale nie do końca wiedzą jak (11 osób),
- wiedzą, że wykres powinien wyglądać zupełnie inaczej, ale nie umieją go wykonać w arkuszu kalkulacyjnym (2 osoby),
- na wykresie brakuje jedynie kilku elementów, ale nie umieją ich dodać lub nie umieją zmodyfikować wyglądu niektórych elementów (5 osób); głównie sygnalizowano nieumiejętność dodania legendy, tytułu wykresu i etykiet danych.

Jedna osoba napisała, że nie dokończyła wykresu, ale nie podała z jakiego powodu.

Deklarowany średni czas, jaki studenci poświęcili na utworzenie wykresu, w tej grupie był nieco wyższy i wyniósł: 9,1 min. Maksymalny czas, jaki jeden ze studentów spędził nad wykresem to 30 min, a najkrótszy to 3 min. Stopień trudności przygotowania wizualizacji, tych 19 studentów oceniło raczej na średni lub trudny, a własne umiejętności tworzenia wykresów jako zerowe (1 osoba), małe/niewielkie (9 osób) i średnie (również 9 osób).

Co ciekawe, nikt z badanych (zarówno z grupy zadowolonej z utworzonej wizualizacji, jak i tych, którzy chcieliby, aby ich wykres wyglądał inaczej) nie uznał, że przygotowanie wykresu było bardzo trudne (tab. 4.2).

Tabela 4.2. Ocena stopnia trudności przygotowania wizualizacji przez zadowolonych i niezadowolonych z opracowanej przez siebie wizualizacji

Stopień trudności przygotowania wizualizacji	Ocena stopnia trudności przygotowania wizualizacji przez	
	zadowolonych z opracowanej przez siebie wizualizacji	niezadowolonych z opracowanej przez siebie wizualizacji
bardzo łatwy	45%	11%
łatwy	33%	21%
średni	15%	42%
trudny	8%	26%
bardzo trudny	0%	0%

Źródło: opracowanie własne

Zgodne z oczekiwaniami wydaje się zestawienie deklarowanego czasu poświęconego na przygotowanie wykresu z oceną własnych umiejętności w tym zakresie (tab. 4.3). Im niższe umiejętności, tym dłuższy czas.

Tabela 4.3. Zestawienie oceny własnych umiejętności tworzenia wykresów z czasem poświęconym na przygotowanie wizualizacji

Subiektywna ocena własnych umiejętności tworzenia wykresów (w nawiasie liczba wskazań)	Czas deklarowany przez badanego poświęcony na przygotowanie wizualizacji (uśredniony, w minutach)
zerowe lub prawie zerowe (2)	10
niewielkie/male (31)	9
średnie (56)	6,8
duże/spore (10)	3,5

Źródło: opracowanie własne.

Ostatnie pytanie ankiety dotyczyło odczuwania przez studentów potrzeby poszerzenia swojej wiedzy w zakresie tworzenia wykresów. 46 osób uznało, że tak, chciałoby uzupełnić i poszerzyć umiejętności w tym zakresie, by móc dobrze i szybko tworzyć wykresy. W tej grupie było dwóch studentów, którzy swoje obecne umiejętności oceniają jako spore. Tym bardziej cieszy to, że chcieliby je dalej pogłębiać. 53 badanych wystarczy to, co umieją, nie odczuwają potrzeby dalszego rozwoju w tym obszarze. Czworo z nich uznało, że nie sądzą, aby było im to kiedykolwiek przydatne.

Krótką charakterystyką osób, które przygotowały w 100% poprawne wizualizacje, wygląda następująco:

- średnio zajęło im to 6,1 min,
- w grupie tej było 7 studentek i 3 studentów,
- 4 osoby to studenci studiów pierwszego stopnia, a 6 studiów magisterskich,
- 3 osoby z tej grupy swoje umiejętności tworzenia wykresów oceniają jako duże, 6 osób jako średnie i jedna jako małe,
- 5 osób odczuwa potrzebę dalszego doskonalenia w tym zakresie i tyle samo uważa, że wystarczy im to, co umieją.

Poniżej zbadano występowanie potencjalnych zależności między sporządzeniem poprawnych wizualizacji a płcią oraz czasem studiowania. Można bowiem oczekiwać, że studenci studiów II stopnia (jako starsi, z większą wiedzą i doświadczeniem) w stosunkowo większej liczbie przypadków sporządzą poprawne wizualizacje niż studenci studiów I stopnia.

Istnienie takiej zależności zbadano testem niezależności χ^2 . Dane empiryczne umieszczone zostały w tablicy czteropolowej o w wierszach i k kolumnach, gdzie n_i i n_j oznaczają liczebności brzegowe (tab. 4.4).

Tabela 4.4. Poprawność wizualizacji a stopień studiów

Stopień studiów	Poprawne wizualizacje		n_i
	tak	nie	
I (licencjacki)	4	63	67
II (magisterski)	6	26	32
n_j	10	89	99

Źródło: opracowanie własne.

Dla przyjętego poziomu istotności $\alpha = 0,05$ i dla 1 stopnia swobody wartość krytyczna χ^2_{α} w tablicy rozkładu χ^2 wynosi 3,841.

Wartość statystyki χ^2 obliczono według wzoru 4.1.

$$\chi^2 = \sum_{i=1}^w \sum_{j=1}^k \frac{(n_{ij} - E_{ij})^2}{E_{ij}} \quad (4.1),$$

gdzie:

E_{ij} – liczebności teoretyczne (oczekiwane), wyliczone według formuły:

$$E_{ij} = \frac{(n_i * n_j)}{n} \quad (4.2),$$

gdzie:

n – liczebność próby.

Statystyka χ^2 wynosi 3,895. Ponieważ $\chi^2 > \chi^2_\alpha$ to są podstawy do twierdzenia, że studenci studiów magisterskich są lepiej przygotowani w zakresie tworzenia wizualizacji niż studenci studiów licencjackich. Wyniki są istotne statystycznie dla poziomu istotności $\alpha = 0,05$, czyli przy założeniu 5% szans popełnienia błędu przy wnioskowaniu.

Co ciekawe, ani jeden student z kierunku IwB (Informatyka w biznesie) nie wykonał w pełni poprawnej wizualizacji. A można by oczekiwać, że ta grupa ma większe umiejętności (przynajmniej ze względu na narzędzia, tj. znajomość i obsługę oprogramowania) niż studenci kierunków „nieinformatycznych”.

Testem niezależności c χ^2 zbadano również istnienie zależności między poprawnością przygotowanej wizualizacji a płcią jej autora (tab. 4.5).

Tabela 4.5. Poprawność wizualizacji a płeć jej autora

Płeć autora wizualizacji	Poprawne wizualizacje		n_i
	tak	nie	
kobieta	7	42	49
mężczyzna	3	47	50
n_j	10	89	99

Źródło: opracowanie własne.

Wartość krytyczna χ^2_α w tablicy rozkładu χ^2 dla 1 stopnia swobody i $\alpha = 0,05$ wynosi 3,841, a obliczona według wzoru (4.1) wartość statystyki χ^2 wyniosła 1,871. Przeprowadzony test χ^2 nie wykazał zatem, jakoby jedna z grup (kobiety lub mężczyźni) przygotowywała lepsze wizualizacje.

4.5. Podsumowanie wyników badań

W niniejszym rozdziale zaprezentowano wyniki badań osób pretendujących w przyszłości do stanowisk menedżerskich, pod kątem ich umiejętności w zakresie wizualizacji. Żyjemy w czasach przepełnionych danymi i informacjami. Formy graficzne coraz częściej są główną podstawą wnioskowania i podejmowania ważnych decyzji. W tym kontekście uzyskane wyniki nie napawają optymizmem. Ogólny

poziom umiejętności badanych w zakresie tworzenia wizualizacji ocenić można jako bardzo niski – jedynie 10% wykonanych wizualizacji było poprawnych.

Zbadano też zależność między umiejętnością sporządzenia poprawnych wizualizacji a czasem studiowania badanych oraz ich płcią. W pierwszym przypadku dane wskazały, że studenci studiów magisterskich są lepiej przygotowani do tworzenia wizualizacji niż studenci studiów licencjackich, natomiast w drugim – nie zaobserwowano korelacji. Zamierzeniem autorki było zwrócenie uwagi na pilną potrzebę kształcenia zarówno w zakresie umiejętności merytorycznych związanych z tworzeniem poprawnych wizualizacji, jak i umiejętności narzędziowych, dotyczących obsługi konkretnych programów informatycznych. Nie ulega bowiem wątpliwości, że rozwój technologii znacznie ułatwia tworzenie efektywnych wizualizacji, a także ich dystrybucję, dzięki taniej formie cyfrowej.

I podobnie jak w podsumowaniu w rozdziale trzecim, po badaniu nasuwa się ten sam wniosek, że konieczne jest podnoszenie poziomu świadomości społeczeństwa w zakresie jakości wizualizacji, z jakimi na co dzień mamy do czynienia i tym samym zwiększenie uwagi przy ich interpretacji.

Jeśli wizualizacje w raportach i sprawozdaniach dla menedżerów przygotowywane są w taki sposób, jak w ramach zaprezentowanego badania, to decyzje i działania podejmowane na ich podstawie mogą być bardzo ryzykowne biznesowo. Chaos informacyjny, błędna lub niepełna informacja, niewychwycona w odpowiednim czasie, może przyczyniać się do znacznych strat finansowych. Wizualizacja ma ułatwiać i przyspieszać percepcję informacji, a z badań wynika, że przy obecnym, niskim poziomie umiejętności jej tworzenia, może stanowić źródło problemów.

Mimo że badanie nie jest reprezentatywne i nie pozwala na uogólnianie wniosków, to ma charakter poznawczy, zwraca uwagę na problem, którego rozmiary nie są znane i otwiera przestrzeń dla innych eksperymentów i analiz poziomu umiejętności społeczeństwa (lub wybranych grup) w zakresie tworzenia poprawnych wizualizacji.

Rezultaty badania skłaniają do refleksji nad częstotliwością, z jaką, całkowicie nieświadomie, możemy mieć do czynienia z błędnym lub niepełnym przekazem. Warto zatem zwiększyć czujność i bardziej krytycznie przyglądać się wszelkim formom graficznym prezentującym ważne dla nas treści.

Zakończenie

W dzisiejszym świecie coraz częściej słowo zastępowane jest obrazem. Wynika to z prozaicznej przyczyny – nadmiaru informacji oraz ograniczonych możliwości i zasobów czasowych do ich przetworzenia. By zatem usprawnić i zwiększyć efektywność komunikacji, ważne jest zrozumienie celu i zasad tworzenia przekazów wizualnych. Znaczenia nabiera właściwe przygotowanie projektanta grafiki w zakresie eksploracji danych, analizy kontekstu, określania funkcji przekazu, doboru odpowiedniej formy oraz dostosowania się do adresata. Twórca wizualizacji ma bardzo duży wpływ na sposób jej odbioru. Jego zadaniem jest wyodrębnienie najistotniejszych treści i związków, które chce przekazać. Oznacza to, że musi najpierw dobrze poznać i zrozumieć dane/informacje/wiedzę, musi dotrzeć do meritum treści i jej kontekstu, tak, by odpowiednio je zestawić, pokazać główny sens i kluczowe zależności. Tylko wtedy tworzone przez niego wizualizacje będą miały dodaną wartość poznawczą i będą istotne dla lepszej percepcji treści przez odbiorców.

Korzystający z wizualizacji powinni mieć świadomość nieuniknionego skumulowania informacji, ale też oddziaływania na ich emocje i zachować pewną ostrożność, zwłaszcza, jeśli na bazie przekazu graficznego zamierzają podejmować ważne decyzje, na przykład biznesowe. Dobrze przygotowana wizualizacja może ułatwić dostrzeżenie zależności lub odkrycie tendencji. Jest też świetnym sposobem na zapamiętanie faktów. Przyspiesza komunikację i poprawia transfer wiedzy. Każdy odbiorca wizualizacji powinien być jednak świadomy możliwych zniekształceń i celowych lub przypadkowych manipulacji, jakie może zastosować projektant, na przykład przez prezentację wybiórczych danych lub niewłaściwy dobór formy wizualizacji. Prezentacja i analiza różnych form deformacji wykresów miała na celu zwiększenie koncentracji czytelników podczas odczytu wizualizacji. Może pozwolić to ustrzec się przed błędną interpretacją faktów oraz podjęciem niewłaściwych decyzji, co jest niebagatelne zwłaszcza w sferze biznesu.

Przedstawione w ostatnim rozdziale książki wyniki badań ukazują duże braki w umiejętnościach tworzenia wizualizacji przez młode, wykształcone osoby wkraczające na rynek pracy. Brak kompetencji widoczny jest w szerokim zakresie: poczynając od doboru danych, poprzez wybór formy wizualizacji, kończąc na aspektach edycyjnych, decydujących o czytelności i dostępności grafiki.

Autorka odsyła czytelników do rosnącej liczby pozycji literaturowych, serwisów internetowych, rozwiązań w postaci selektorów wykresów i obecnie niezastąpionych już narzędzi informatycznych do tworzenia wizualizacji. Poleca też szeroki wachlarz szkoleń i warsztatów oferowanych online przez specjalistów w tym zakresie.

Bibliografia

- Bauer M., Johnson-Laird P. (1993). How diagrams can improve reasoning. *Psychological Science* 4(6), 372–378. DOI: 10.1111/j.1467-9280.1993.tb00584.x.
- Bertschi S., Bresciani S., Crawford T., Goebel R., Kienreich W., Lindner M., Sabol V., Moere A. (2011). What is knowledge visualization? Perspectives on an emerging discipline, w: *Proceedings of the International Conference on Information Visualisation*, 329–336. DOI: 10.1109/IV.2011.58.
- Biecek P. (2014). *Odkrywać! Ujawniać! Objaśniać! Zbiór esejów o sztuce pokazywania danych*. Warszawa: Fundacja Naukowa SmarterPoland.pl. <http://www.biecek.pl/Eseje/>.
- Bresciani S., Eppler M.J. (2015). The Pitfalls of Visual Representations. *SAGE Open* 5(4). DOI: 10.1177/2158244015611451.
- Burkhard R.A. (2004). Learning from architects: the difference between knowledge visualization and information visualization, w: *Proceedings of Eighth International Conference on Information Visualization* (04.2004), 519–524. IEEE, London
- Burkhard R.A. (2005). Towards a Framework and a Model for Knowledge Visualization: Synergies Between Information and Knowledge Visualization. Conference *Knowledge and Information Visualization Searching for Synergies*. DOI: 10.1007/11510154_13.
- Cairo A. (2012). *The Functional Art: An introduction to information graphics and visualization*. Berkeley: New Riders.
- Cairo A. (2016). *The Truthful Art: Data, Charts, and Maps for Communication*. Berkeley: New Riders.
- Cairo A. (2018). *In visualization, white space is your friend*. <https://questionsindataviz.com/2018/01/06/is-white-space-always-your-friend/>.
- Cairo A. (2020). *How Charts Lie: Getting Smarter about Visual Information*. New York: W.W. Norton & Company.

- Camões J. (2016). *Data at Work: Best practices for creating effective charts and information graphics in Microsoft Excel*. Berkeley: New Riders.
- Card S.K., Mackinlay J.D., Shneiderman B. (1999). *Readings in Information Visualization; Using Vision to think*. Los Altos, CA: Morgan Kaufmann.
- Chan V. (2017) *Getting it right: why infographics are not the same as data visualizations*. <https://blog.prototypr.io/getting-it-right-why-infographics-are-not-the-same-as-data-visualizations-a23da7de745e>.
- Chen M., Ebert D., Hagen H., Laramée R.S., Van Liere R., Ma K.-L., Ribarsky W., Scheuermann, G., Silver D. (2009). Data, Information, and Knowledge in Visualization. *Computer Graphics and Applications, IEEE*. 29(1), 12–19. DOI: 10.1109/MCG.2009.6.
- Cicharski B. (2021). Kłamstwo ma krótką oś Y, czyli jak wykreślić nieprawdę. *Finanse i Controlling* 73.
- Czapiewski B. (2010). *Projektowanie wykresu krok po kroku*. <https://skuteczneraporty.pl/blog/projektowanie-wykresy-krok-po-kroku/>.
- Dylak S. (2012). *Alfabetyzacja wizualna jako kompetencja współczesnego człowieka*, w: W. Skrzydlewski, S. Dylak (red.), *Media – Edukacja – Kultura*. Poznań–Rzeszów: Polskie Towarzystwo Technologii i Mediów Edukacyjnych, 119–132.
- Ellis W.D. (1938). *A Source Book of Gestalt Psychology*. New York: Harcourt, Brace & World. DOI: 10.1037/11496-000.
- Emery A.A. (2021). *Designing accessible data visualizations. 10 Quick Wins*. <https://depictdatastudio.gumroad.com/l/accessibility>, <https://gumroad.com/l/NILdp>.
- Eppler M.J. (2011). What is an effective knowledge visualization? Insights from a review of seminal concepts, w: *Proceedings of the International Conference on information Visualisation*, 3–12. DOI: 10.1109/IV.2011.13.
- Falkowitz R. (2019). *Information Visualization or Data Visualization?* Concentric Circle Consulting. <https://www.3cs.ch/information-visualization-data-visualization/>.
- Few S. (2009). *Now You See It: Simple Visualization Techniques for Quantitative Analysis*. Oakland: Analytics Press.
- Few S. (2012). *Show Me the Numbers: Designing Tables and Graphs to Enlighten*. Oakland: Analytics Press.
- Few S. (2013). *Information Dashboard Design: Displaying Data for At-a-glance Monitoring*. Oakland: Analytics Press.
- Franconeri S. (2019). *Multiple views on how to choose a visualization*. <https://medium.com/multiple-views-visualization-research-explained/multiple-views-on-how-to-choose-a-visualization-b3ffc99fcddc>.

- Gatto M.A.C. (2015). *Making Research Useful: Current Challenges and Good Practices in Data Visualization*. Oxford: Reuters Institute for the Study of Journalism, University of Oxford.
- Glenberg A., Langston M. (1992). Comprehension of illustrated text: pictures help to build mental models. *Journal of Memory and Language* 31(2), 129–151. DOI: 10.1016/0749-596X(92)90008-L.
- Kahneman D. (2012). *Pułapki myślenia. O myśleniu szybkim i wolnym*. Poznań: Media Rodzina.
- Kelleher C., Wagener T. (2011). Ten guidelines for effective data visualization in scientific publications. *Environmental Modelling and Software* 26(6), 822–827. DOI: 10.1016/j.envsoft.2010.12.006.
- Kirk A. (2019). *Data Visualisation: A Handbook for Data Driven Design*. Thousand Oaks: SAGE Publications Ltd.
- Koffka K. (1935). *The Principles of Gestalt Psychology*. New York: Harcourt, Brace & World.
- Kosslyn S.M. (1980). *Images and Mind*. Cambridge: Harvard University Press.
- Krok E. (2013). *Determinanty skłonności pracowników do dzielenia się wiedzą – koncepcja diagnozy dla potrzeb zarządzania*. Szczecin: Volumina.
- Krok E. (2021) Visualization on charts – manipulations and distortions. Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 25th International Conference KES2021. *Procedia Computer Science* 192, 3932–3944. DOI: 10.1016/j.procs.2021.09.168.
- Krok E. (2022). Assessment of students' skills in visual communication of information. Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 26th International Conference KES2022. *Procedia Computer Science* 207, 2891–2900. DOI: 10.1016/j.procs.2022.09.347.
- Król K. (2019). Map charts: visualisation of statistical data on a background map – case study. *Geomatics, Landmanagement and Landscape (GLL)* 4, 171–181. DOI: 10.15576/GLL/2019.4.171.
- Larkin J., Simon H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science* 11, 65–99. DOI: 10.1111/j.1551-6708.1987.tb00863.x.
- Leo S. (2019). Mistakes, we've drawn a few. Learning from our errors in data visualisation. *The Economist*. <https://medium.economist.com/mistakes-weve-drawn-a-few-8cdd8a42d368>.
- Manovich L. (2010). What is visualization? *The Journal of the Initiative for Digital Humanities, Media, and Culture* 2(1), 1–32.

- Marchese F.T., Banissi E. (2013). *Knowledge Visualization Currents: From Text to Art to Culture*. London: Springer-Verlag. DOI: 10.1007/978-1-4471-4303-1.
- Medyńska-Gulij B. (2011). *Kartografia i geowizualizacja*. Warszawa: Wydawnictwo Naukowe PWN.
- Meirelles I. (2013). *Design for Information: An Introduction to the Histories, Theories, and Best Practices Behind Effective Information Visualizations*. Beverly: Rockport Publishers.
- Meyer R. (2010). Knowledge Visualization. *Trends in Information Visualization* 32, 63–84.
- Midway S.R. (2020). Principles of Effective Data Visualization. *Patterns Perspective* 1(9). DOI: 10.1016/j.patter.2020.100141.
- Miller G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63, 81–97. DOI: 10.1037/h0043158.
- Mills J.H., Thurlow A., Mills A.J. (2010). Making sense of sensemaking: the critical sensemaking approach. *Qualitative Research in Organizations and Management: An International Journal* 5(2), 182–195. DOI: 10.1108/17465641011068857.
- Motylińska P. (2014). Wybrane zagadnienia wizualizacji w nauce o informacji, w: A. Korycińska-Huras, M. Janiak (red.), *Komunikacja naukowa w środowisku cyfrowym: badania, zasoby, użytkownicy*. Warszawa: Wydawnictwo SBP, 406–427.
- Munzner T. (2009). Visualization. w: P. Shirley, S. Marschner (red.), *Fundamentals of Computer Graphics*. Natick: AK Peters, 675–707. DOI: 10.1201/9781315372198-26.
- Nielsen C.B. (2016). Visualization: A Mind-Machine Interface for Discovery. *Trends in Genetics* 32(2), 73–75. DOI: 10.1016/j.tig.2015.12.002.
- Nonaka I. (1991). The Knowledge-Creating Company. *Harvard Business Review* 69(6), 96–104.
- Novick L.R. (2001). Spatial diagrams: Key instruments in the toolbox for thought. *The psychology of learning and motivation* 40, 279–325. DOI: 10.1016/S0079-7421(00)80023-7.
- Nussbaumer Knaflitz C. (2015). *Storytelling with Data: A Data Visualization Guide for Business Professionals*. Hoboken: Wiley.
- Olshannikova E., Ometov A., Koucheryavy Y., Olsson T. (2015). Visualizing Big Data with augmented and virtual reality: challenges and research agenda. *Journal of Big Data* 2(1), 22. DOI: 10.1186/s40537-015-0031-2.

- Pugsley S. (2022). *Telling a Great Data Story: A Visualization Decision Tree*. <https://www.kdnuggets.com/2021/02/telling-great-data-story-visualization-decision-tree.html>.
- Renaud K., van Biljon, J. (2017). Charting the Path towards Effective Knowledge Visualisations, w: *Proceedings of the Southern African Institute of Computer Scientists and Information Technologists, South Africa, (SAICSIT'17)*. Association for Computing Machinery. DOI: 10.1145/3129416.3129421.
- Riche N.H., Hurter Ch., Diakopoulos N., Carpendale S. (2018). *Data-Driven Storytelling*. Boca Raton: A K Peters/CRC Press.
- Robbins N.B. (2013). *Creating More Effective Graphs*. Katy: Chart House.
- Roberts J.C., Ritsos P.D., Badam S.K., Brodbeck D., Kennedy J., Elmqvist N. (2014). Visualization beyond the desktop-the next big thing. *IEEE Computer Graphics and Applications* 34(6), 26–34. DOI: 10.1109/MCG.2014.82.
- Schwabish J.A. (2016). *Better Presentations: A Guide for Scholars, Researchers, and Wonks*. New York: Columbia University Press. DOI: 10.7312/schw17520.
- Schwabish J.A. (2020). *Elevate the Debate: A Multilayered Approach to Communicating Your Research*. Hoboken: Wiley.
- Shepard R.N., Cooper L.A. (1982). *Mental Images and Their Transformations*. Cambridge: MIT Press.
- Shirley P., Marschner S. (red.) (2009). *Fundamentals of Computer Graphics*. Natick: AK Peters.
- Stanimir A. (2012). Różne techniki prezentacji powiązań kategorii zmiennych niemytrycznych. *Ekonometria* 3(37), 9–25.
- Steczkowski J., Zeliaś A. (1981). *Statystyczne metody analizy cech jakościowych*. Warszawa: PWE.
- Stevens S.S. (1959). Measurement, psychophysics and utility, w: C.W. Churchman, P. Ratoosh (red.), *Measurement: Definitions and Theories*. New York: Wiley.
- Szews P. (2017). *Wizualizacja przekazu w dziennikarstwie. Infografika i data journalism w wybranych mediach*. Praca doktorska, Uniwersytet Łódzki.
- Taei P. (2018). *What is an infographic? And how is it different from a data visualization?* <https://towardsdatascience.com/what-is-an-infographic-and-how-is-it-different-from-a-data-visualization-a92c23b35197>.
- Tal A., Wansink B. (2016). Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy. *Public Understanding of Science* 25(1), 117–125. DOI: 10.1177/0963662514549688.
- Tufte E.R. (2007). *The Visual Display of Quantitative Information*. Cheshire: Graphics Press.

- Valkanova N., Jorda S., Vande Moere A. (2015). Public visualization displays of citizen data: Design, impact and implications. *International Journal of Human Computer Studies* 81, 4–16. DOI: 10.1016/j.ijhcs.2015.02.005.
- Van Biljon J., Renaud K. (2015). Do Visualizations Ease Dissertation Assessment?, w: *Proceedings of the 44th Annual Southern African Computer Lecturers Association*, 177–185.
- Van Stralen D. (2015). Ambiguity. *Journal of Contingencies and Crisis Management* 23(2), 47–53. DOI: 10.1111/1468-5973.12082.
- Ware C. (2013). *Information Visualization. Perception for Design*. Burlington: Morgan Kaufmann.
- Wexler S., Shaffer J., Cotgreave A. (2017). *The Big Book of Dashboards: Visualizing Your Data Using Real-World Business Scenarios*. Hoboken: Wiley. DOI: 10.1002/9781119283089.
- Wilke C.O. (2020). *Podstawy wizualizacji danych. Zasady tworzenia atrakcyjnych wykresów*. Gliwice: Helion.
- Wong Dona M. (2013). *The Wall Street Journal Guide to Information Graphics: The Dos and Don'ts of Presenting Data, Facts, and Figures*. New York: W.W. Norton & Company.
- Yaacob S., Liang H., Mohamad N. (2017). Towards a characterization of complex visualizations. *Open International Journal of Informatics (OIJI)* 5, 19–29.
- Yau N. (2011). *Visualize This: The Flowing Data Guide to Design, Visualization, and Statistics*. Hoboken: Wiley.
- Yusoff Z., Katmon S.A., Ahmad M.N., Miswan S.H.M. (2013). Visual representation: Enhancing students's learning engagement through knowledge visualization, w: *International Conference on Informatics and Creative Multimedia*, 242–247. DOI: 10.1109/ICICM.2013.48.
- Zhang Y., He X., Xie J., Wang Z. (2008). Study on the knowledge visualization and creation supported Kmap platform, w: *1st International Workshop on Knowledge Discovery and Data Mining, WKDD*, 154–159.

Portale i strony internetowe, do których odwołania pojawiają się w tekście

<http://chartchoosercards.com/>

<http://chartmaker.visualisingdata.com/>

<http://visualizationuniverse.com/>

<https://adioma.com/>

<https://coolinfographics.com/>

<https://datavizcatalogue.com/>

<https://datavizproject.com/>
<https://depictdatastudio.com/>
https://experception.net/_
<https://extremepresentation.com/>
<https://flowingdata.com/>
<https://gravyanecdote.com>
<https://infonewt.com/>
<https://insights.datylon.com/stories>
[https://kulturawrazliwa.pl/wiedza/
jak-stworzyc-tekst-alternatywny-krotki-poradnik/](https://kulturawrazliwa.pl/wiedza/jak-stworzyc-tekst-alternatywny-krotki-poradnik/)
<https://policyviz.com/>
<https://readable.com/>
<https://stephanieevergreen.com/>
[https://support.microsoft.com/pl-pl/office/dost%C4%99pne-typy-
wykres%C3%B3w-w-pakiecie-office-a6187218-807e-4103-9e0a-27cdb19afb90](https://support.microsoft.com/pl-pl/office/dost%C4%99pne-typy-wykres%C3%B3w-w-pakiecie-office-a6187218-807e-4103-9e0a-27cdb19afb90)
<https://thenounproject.com/>
<https://www.data-to-viz.com/>
<https://www.datawrapper.de/>
<https://www.datylon.com/>
https://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=00002o
<https://www.esri.pl/arcgis/>
<https://www.ferdio.com/>
<https://www.geckoboard.com/>
<https://www.iconfinder.com/>
<https://www.juiceanalytics.com/chartchooser/>
[https://www.kdnuggets.com/wp-content/uploads/pugsley-visualization-decision-
tree-kd-v12.jpg](https://www.kdnuggets.com/wp-content/uploads/pugsley-visualization-decision-tree-kd-v12.jpg)
<https://www.qlik.com/blog/patrik-lundblad/>
<https://www.statsoft.pl/textbook/>
<https://www.visualisingdata.com/>
<https://www.visual-literacy.org/>

Spis rysunków

Rysunek 1.1. Elementy, które należy przeanalizować przed rozpoczęciem wizualizacji	10
Rysunek 1.2. Zestaw schematów w pakiecie MS Office: lista, proces, cykl, hierarchia, relacja, macierz, piramida	19
Rysunek 1.3. Przykładowe rodzaje wykresów w pakiecie MS Office: kolumnowy, słupkowy, liniowy, warstwowy, histogram, kołowy, punktowy	20
Rysunek 1.4. Przykładowe rodzaje map tematycznych: sygnaturowa, zasięgu, przepływu, termiczna, kartogram, kartodiagram	22
Rysunek 2.1. Wskazówki dotyczące wizualizacji danych od firmy Geckoboard – plakat	30
Rysunek 2.2. Wykres przed modyfikacjami zwiększającymi czytelność	35
Rysunek 2.3. Wykres po modyfikacjach zwiększających czytelność	35
Rysunek 2.4. Przekształcenia liczb, prowadzące do zmniejszenia wysiłku odbiorcy związanego z liczeniem i tym samym do zwiększenia czytelności wizualizacji	36
Rysunek 2.5. Układ okresowy metod wizualizacji opracowany w ramach projektu Visual Literacy	39
Rysunek 2.6. „Schody do wizualnej doskonałości” – projekt Visual Literacy	40
Rysunek 2.7. Strona główna projektu „Wszechświat wizualizacji” prowadzonego przez portal Adioma i Google News Lab	42
Rysunek 2.8. Narzędzia specjalistyczne do wizualizacji danych analizowane w projekcie „Wszechświat wizualizacji” – sortowanie według zainteresowania wyszukiwaniem	43

Rysunek 2.9. Narzędzia ogólne do wizualizacji danych analizowane w projekcie „Wszechświat wizualizacji” – sortowanie według zainteresowania wyszukiwaniem	44
Rysunek 2.10. Katalog wizualizacji danych opracowany przez Severino Ribecce.....	44
Rysunek 2.11. Szczegółowe informacje o wykresie kołowym – przykład z katalogu wizualizacji danych Severino Ribecci	45
Rysunek 2.12. Podział wykresów na kategorie w katalogu wizualizacji danych Severino Ribecci.....	47
Rysunek 2.13. Przewodnik po wyborze wykresów opracowany przez Andrew Abele	48
Rysunek 2.14. Selektor wykresów online opracowany przez zespół z Juiceanalytics.....	50
Rysunek 2.15. Fragment witryny projektu Data Viz	51
Rysunek 2.16. Fragment zrzutu ekranu dotyczący informacji o wykresie liniowym – projekt Data Viz	52
Rysunek 2.17. Fragment zrzutu ekranu dotyczący informacji o wykresie kołowym – projekt Data Viz.....	52
Rysunek 2.18. Fragment zrzutu ekranu dotyczący informacji o wykresie słupkowym z witryny Depict Data Studio.....	53
Rysunek 2.19. Fragment zrzutu ekranu z katalogiem kreatora wykresów (Chartmaker Directory), opracowanym przez Andy’ego Kirka.....	54
Rysunek 2.20. Fragment zrzutu ekranu pokazujący możliwości współtworzenia witryny z katalogiem kreatora wykresów, opracowanym przez Andy’ego Kirka.....	55
Rysunek 2.21. Selektor rekordów w postaci drzewa decyzyjnego – propozycja Stana Pugsleya	56
Rysunek 2.22. Jaki wykres wybrać – krótki poradnik Steve’a Franconeriego.....	57
Rysunek 2.23. Drzewa decyzyjne ułatwiające dobór wykresu dla różnego rodzaju danych wejściowych.....	58
Rysunek 2.24. Funkcjonalności oferowane na platformie https://www.datylon.com/	60
Rysunek 2.25. Kategorie ikon i przykładowe ikony w ramach kategorii, dostępne na portalu Adioma	61
Rysunek 2.26. Kategorie infografik zaproponowane przez Annę Vital	62

Rysunek 2.27. Przykładowa infografika przedstawiająca jak zwizualizować artykuł.....	63
Rysunek 4.1. Mało przejrzyste i nieadekwatne do prezentowanej treści wykresy kołowe z błędną legendą	93
Rysunek 4.2. Nieprawidłowo dobrany do prezentowanych treści wykres lejkowy.....	93
Rysunek 4.3. Wykres kolumnowy – dwie serie danych dla poszczególnych miesięcy.....	94
Rysunek 4.4. Wykres kolumnowy – sześć serii danych dla dwóch lat.....	94
Rysunek 4.5. Wykres słupkowy – jedna seria danych	95
Rysunek 4.6. Niepoprawnie sporządzone wykresy kolumnowe – niewnosząca żadnych treści legenda	96
Rysunek 4.7. Brak informacji o miesiącach, w jakich zatrudniano pracowników	97
Rysunek 4.8. Błędnie przedstawiony okres zatrudniania – 12 miesięcy jednego roku.....	98
Rysunek 4.9. Błędnie przedstawiony okres zatrudniania, błędny tytuł i legenda wykresu	98
Rysunek 4.10. Błędnie przedstawiony okres zatrudniania – sześć ostatnich miesięcy 2019 roku i sześć pierwszych kolejnego roku.....	99
Rysunek 4.11. Brak nazw miesięcy, w których następowało zatrudnianie oraz brak informacji, że wykres przedstawia liczbę nowo zatrudnionych pracowników (a nie pracowników ogółem).....	99
Rysunek 4.12. Brak informacji co zaprezentowano na wykresie – przykład 1	100
Rysunek 4.13. Brak informacji co zaprezentowano na wykresie – przykład 2	101
Rysunek 4.14. Brak informacji co zaprezentowano na wykresie – przykład 3	101
Rysunek 4.15. Pominięcie niektórych danych – brak danych dla czerwca 2020 roku.....	102
Rysunek 4.16. Niechronologiczny układ prezentacji danych utrudniający zrozumienie przekazywanych treści	102
Rysunek 4.17. Brak jakiegokolwiek związku wykresu z informacją, jaką miał prezentować.....	103

Spis tabel

Tabela 1.1. Dane prezentujące wyposażenie klubu seniora	15
Tabela 3.1. Dane dotyczące sprzedaży	78
Tabela 4.1. Formy wizualizacji wybrane przez studentów.....	92
Tabela 4.2. Ocena stopnia trudności przygotowania wizualizacji przez zadowolonych i niezadowolonych z opracowanej przez siebie wizualizacji	105
Tabela 4.3. Zestawienie oceny własnych umiejętności tworzenia wykresów z czasem poświęconym na przygotowanie wizualizacji.....	105
Tabela 4.4. Poprawność wizualizacji a stopień studiów	106
Tabela 4.5. Poprawność wizualizacji a płeć jej autora	107

Spis wykresów

Wykres 1.1. Wizualizacja zgodna z potrzebami organizatora spotkań dla seniorów.....	16
Wykres 1.2. Wizualizacja zgodna z potrzebami pracownika działu finansowego klubu seniora – sortowanie według wartości.....	17
Wykres 1.3. Wizualizacja zgodna z potrzebami pracownika odpowiedzialnego za wyposażenie klubu seniora.....	17
Wykres 3.1. Przychody ze sprzedaży.....	66
Wykres 3.2. Deformacja wykresu 3.1 – przykład 1.....	67
Wykres 3.3. Deformacja wykresu 3.1 – przykład 2.....	68
Wykres 3.4. Deformacja wykresu 3.1 – przykład 3.....	68
Wykres 3.5. Porównanie średnich zarobków w zawodach A i B.....	69
Wykres 3.6. Deformacja wykresu 3.5 – przykład 1.....	69
Wykres 3.7. Deformacja wykresu 3.5 – przykład 2.....	69
Wykres 3.8. Zasady podziału nagród w konkursie.....	71
Wykres 3.9. Deformacja wykresu 3.8.....	72
Wykres 3.10. Liczba zwolnień w firmie A.....	72
Wykres 3.11. Liczba zwolnień w firmie B.....	73
Wykres 3.12. Udział w rynku.....	74
Wykres 3.13. Deformacja wykresu 3.12.....	74
Wykres 3.14. Liczba rekrutowanych studentów.....	75
Wykres 3.15. Dane dotyczące działalności firmy A.....	76
Wykres 3.16. Dane z wykresu 3.15 skumulowane do 100%.....	76
Wykres 3.17. Liczba wykonanych podłóg w wartościach bezwzględnych.....	77
Wykres 3.18. Liczba wykonanych podłóg w ujęciu procentowym.....	77

Wykres 3.19. Dane dotyczące sprzedaży	79
Wykres 3.20. Zachorowania na chorobę X w Polsce.....	80
Wykres 3.21. Wybiórcze dane z wykresu 3.20	80
Wykres 3.22. Sytuacja Polski na tle krajów UE	81
Wykres 3.23. Wybiórcze dane z wykresu 3.22	81
Wykres 3.24. Średnia oglądalność w ostatnim roku	82
Wykres 3.25. Oglądalność konkretnego dnia.....	82
Wykres 3.26. Porównanie eksportu firm X i Y	83
Wykres 3.27. Wybiórcze dane z wykresu 3.26	83
Wykres 3.28. Liczba projektów zrealizowanych przez organizacje A i B w latach 2018–2020.....	84
Wykres 3.29. Zmiana stopy bezrobocia	85
Wykres 3.30. Deformacja wykresu 3.29.....	85
Wykres 3.31. Liczba klientów w godzinach otwarcia sklepu – przykład deformacji histogramu.....	86
Wykres 3.32. Liczba turystów zwiedzających obiekt X – przykład deformacji histogramu.....	86

Streszczenie

Słowa kluczowe: wizualizacja danych, wizualizacja informacji, zniekształcenia na wykresach, dobre praktyki tworzenia wizualizacji

Monografię poświęcono zagadnieniom wizualizacji. Dokonano systematyzacji pojęć, przeanalizowano istotne elementy, których znajomość rzutuje na jakość tworzonych wizualizacji i tym samym wpływa na skuteczność przekazywania treści, co z biznesowego punktu widzenia często decyduje o trafności podejmowanych działań.

W publikacji zawarto wiele cennych wskazówek, dotyczących zasad opracowywania wizualizacji, pochodzących od światowej klasy specjalistów. Poruszono aspekty, jakie wpływają na czytelność i jednoznaczność przekazu, a także na jego dostępność. Dokonano przeglądu witryn i serwisów internetowych poświęconych wizualizacjom, selektorom wykresów oraz narzędziom informatycznym wspierającym tworzenie grafik.

Przedstawiono też możliwości wypaczania faktów i manipulacji przy przekazie informacji, jakie zdarza się stosować projektantom wykresów w celu wywołania u odbiorcy określonego sposobu myślenia o prezentowanej sytuacji. Przeanalizowano przekłamania wynikające ze świadomego wykorzystania funkcjonalności komputerowych programów graficznych stosowanych do tworzenia wykresów, a także zniekształcenia będące efektem celowego doboru takich, a nie innych danych do wizualizacji.

W ostatnim rozdziale monografii zaprezentowano analizę wyników badania umiejętności studentów w zakresie tworzenia wykresów na podstawie opisu słownego, co pozwoliło na identyfikację najczęściej popełnianych błędów ograniczających czytelność wizualizacji i wpływających na jej jakość.

Abstract

Visualization of data, information and knowledge

Keywords: data visualization, information visualization, distortions in charts, good practices for creating visualizations

The monograph was devoted to the issues of visualization. The concepts were systematized, important elements were analyzed, the knowledge of which affects the quality of the created visualizations and thus affects the effectiveness of content transfer, which from a business point of view often determines the accuracy of the actions taken.

The book contains many valuable tips on the principles of developing visualizations, coming from world-class specialists. Aspects that affect the readability and unambiguity of the message, as well as its accessibility, were discussed. A review of websites devoted to visualizations, chart selectors and IT tools supporting the creation of graphics was made.

It also presents the possibilities of distortion of facts and manipulation in the transfer of information, which are sometimes used by chart designers in order to induce a specific way of thinking about the presented situation in the recipient. Distortions resulting from the conscious use of the functionality of computer graphics programs to create charts, as well as distortions resulting from the deliberate selection of such and not other data for visualization, were analyzed.

The last chapter of the monograph contains an analysis of the results of the study of students' skills in creating charts based on a verbal description, which allowed the identification of the most common mistakes that limit the readability of visualization and affect its quality.

