

9. METODY WIZUALIZACJI DANYCH

Autor: Dario Šebalj

W dzisiejszym świecie opartym na danych, zdolność do skutecznego przekładania złożonych zbiorów danych na przejrzyste, intuicyjne wizualizacje jest koniecznością dla organizacji, które chcą efektywnie wykorzystywać swoje dane. Wizualizacja danych wykracza poza czysto estetyczną reprezentację; jest to podstawowy element analizy biznesowej, który pomaga decydentom identyfikować trendy, wartości odstające i wzorce ukryte w surowych danych. W tym rozdziale przedstawiono różne rodzaje wizualizacji, od prostych wykresów, takich jak wykresy słupkowe i liniowe, po bardziej skomplikowane reprezentacje graficzne, takie jak mapy cieplne i wykresy punktowe. Każdy typ wizualizacji służy różnym celom i jest odpowiedni dla różnych zestawów danych, więc dla analityków danych kluczowe znaczenie ma wybór odpowiedniej wizualizacji, aby skutecznie przekazać zamierzony komunikat.

Wizualizacja danych jest procesem przekształcania informacji w kontekst wizualny, taki jak mapa lub wykres, i jest wykorzystywana do ułatwienia ludzkiemu umysłowi zrozumienia danych i wyciągania z nich wniosków. Głównym celem wizualizacji danych jest ułatwienie identyfikacji wzorców, trendów i wartości odstających w dużych zbiorach danych. Typowe rodzaje wizualizacji danych obejmują wykresy, tabele, mapy i pulpity nawigacyjne (ang. dashboardy) (Brush, 2022; GeeksForGeeks, 2024).

Ze względu na rosnącą popularność projektów big data i analityki danych, wizualizacja jest obecnie ważniejsza niż kiedykolwiek. Firmy coraz częściej wykorzystują uczenie maszynowe do gromadzenia ogromnych ilości danych, których przetwarzanie, zrozumienie i wyjaśnienie może być trudne i powolne. Można ten proces przyspieszyć za pomocą wizualizacji, która również ułatwia interesariuszom i właścicielom firm zrozumienie informacji (Brush, 2022). Przed wyborem metody wizualizacji ważne jest zrozumienie kontekstu wizualizacji.

9.1. Zrozumienie kontekstu sytuacyjnego

Nussbaumer Knaflitz (2015) stwierdza, że zrozumienie i kontekstualizacja jest pierwszym i najważniejszym krokiem przed zaangażowaniem się w techniki wizualizacji danych i metody opowiadania historii. Zrozumienie odbiorców jest kluczowym aspektem kontekstu. Nussbaumer Knaflitz podkreśla znaczenie wiedzy o tym, kim są odbiorcy, jaki jest ich poziom

wiedzy i co jest dla nich ważne. Takie zrozumienie zapewnia, że wizualizacja danych i opowiadanie historii są dostosowane do potrzeb i preferencji odbiorców, dzięki czemu informacje są bardziej istotne i angażujące.

Według IBM (b.d.), ogólne informacje podstawowe pomagają odbiorcom zrozumieć znaczenie konkretnego punktu danych. Na przykład, jeśli wskaźnik otwarć wiadomości e-mail firmy jest poniżej średniej, powinniśmy pokazać, jak wskaźnik otwarć wypada w porównaniu z całą branżą, aby zilustrować, że istnieje problem z tym kanałem marketingowym dla firmy. Odbiorcy muszą zrozumieć, jak obecna wydajność wypada w porównaniu z określonym celem, benchmarkiem lub innymi kluczowymi wskaźnikami wydajności (KPI), aby zmotywować się do podjęcia działań.

Należy więc odpowiedzieć na trzy ważne pytania (Nussbaumer Knaflic, 2015; IBM, b.d.):

- **Kto:** Obejmuje to identyfikację odbiorców i zrozumienie ich perspektywy, aby wiedzieć, jak należy dostosować historię. Takie działanie gwarantuje, że wizualizacja jest skierowana bezpośrednio do odbiorców docelowych, dzięki czemu jest bardziej skuteczna i angażująca. Na przykład, podczas gdy kwartalny raport roczny może wymagać jedynie wysokiego poziomu podsumowania danych, analityk finansowy może potrzebować szczegółowej analizy trendów na przestrzeni kilku lat. Decyzja o złożoności, poziomie szczegółowości i wglądu, który należy podkreślić, zależy od tego, kto będzie patrzył na wizualizację.
- **Co:** Odnosi się to do kluczowego komunikatu lub spostrzeżenia, które ma zostać przekazane odbiorcom. Chodzi o to, aby mieć jasność co do działania lub decyzji, na którą wizualizacja danych ma wpłynąć. Kontekst określa cel wizualizacji. Czy ma ona przekonywać, informować, badać czy potwierdzać? Każdy cel może prowadzić do różnych decyzji dotyczących rodzaju wizualizacji i wyróżnionych punktów danych. Na przykład wizualizacja perswazyjna mająca na celu uzyskanie wsparcia dla nowej inicjatywy może koncentrować się na innych danych niż wizualizacja mająca na celu po prostu poinformowanie o dotychczasowych wynikach.
- **Jak:** Chodzi o wybranie najbardziej odpowiednich i skutecznych środków komunikowania historii lub spostrzeżeń, biorąc pod uwagę medium, format i techniki wizualizacji, które będą najlepiej rezonować z docelowymi odbiorcami. Niektóre rodzaje zestawów danych wymagają również specjalnej wizualizacji. Na przykład wykresy punktowe są właściwe do pokazania relacji między dwiema zmiennymi, a wykresy liniowe są dobrym sposobem na pokazanie danych szeregów czasowych. Wizualizacje powinny pomóc odbiorcom zrozumieć główne przesłanie. Nieprawidłowe

rozmieszczenie wykresów i danych może mieć odwrotny skutek i raczej zmylić niż oświecić odbiorców.

Jeśli chodzi o wizualizację i analizę danych, rozróżnia się wizualizację eksploracyjną i wyjaśniającą. Wizualizacja eksploracyjna motywuje użytkownika do zagłębienia się w dane lub temat w celu dokonania własnych odkryć. Wizualizacja wyjaśniająca wysuwa wyniki na pierwszy plan, przekazując czytelnikowi hipotezę lub argument autora (Schwabish, 2021).

Po zbadaniu znaczenia zrozumienia kontekstu sytuacyjnego, w którym wykorzystywane są wizualizacje danych, jasne jest, że ta podstawowa wiedza określa sposób, w jaki informacje są najlepiej przekazywane i postrzegane przez odbiorców.

Kolejnym krytycznym krokiem jest skuteczne zaangażowanie odbiorców. Następny podrozdział opisuje strategie angażowania i utrzymywania zainteresowania odbiorcy. Obejmuje to wybór elementów, które zwiększają atrakcyjność wizualną i czytelność wizualizacji danych oraz zapewniają, że kluczowe spostrzeżenia nie pozostaną niezauważone. Korzystając z technik, które przyciągają wzrok odbiorcy i podkreślają ważne dane, wizualizacje mogą być czymś więcej niż tylko informacją - mogą być wciągające i przekonujące.

9.2. Metody przyciągania uwagi

Podczas projektowania wizualizacji danych bardzo ważne jest, aby uchwycić i skierować uwagę odbiorców. Interakcja między mechaniką widzenia a zasadami percepcji wzrokowej określa, jak skutecznie wizualizacja przekazuje zamierzony komunikat. Zrozumienie, w jaki sposób ludzkie oko postrzega elementy wizualne, jest pierwszym krokiem w tworzeniu atrakcyjnych wizualizacji.

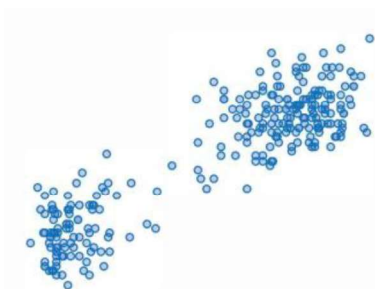


Okolo 70% receptorów zmysłów w naszych ciałach jest przeznaczonych do widzenia (Few, 2012).

Oko jest wstępnie przyciągane do pewnych atrybutów wizualnych, tj. atrybuty te są przetwarzane szybko i automatycznie w systemie wizualnym bez świadomego wysiłku. Atrybuty takie jak kolor, rozmiar, kształt i orientacja mogą być wykorzystane do podkreślenia krytycznych punktów danych lub obszarów w wizualizacji i natychmiast przyciągnąć uwagę odbiorcy.

Według Schwabisha (2021) **teoria Gestalt** opisuje, w jaki sposób ludzie zazwyczaj grupują elementy wizualne. Słowo Gestalt oznacza wzór (Cairo, 2013). Została ona opracowana przez niemieckich psychologów na początku XX wieku. Jeśli chodzi o tworzenie diagramów i innych wizualizacji, sześć zasad teorii Gestalt jest szczególnie pomocnych.

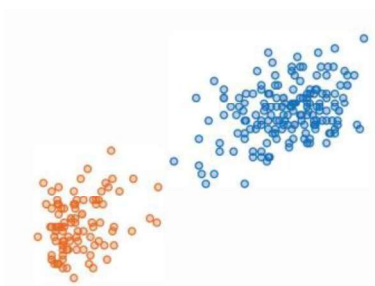
Zasada **bliskości** mówi, że nasza percepcja grupuje obiekty, gdy znajdują się one blisko siebie (np. Rysunek 9.1).



Rysunek 9.1. Bliskość jako zasada teorii Gestalt

Źródło: Schwabish (2021)

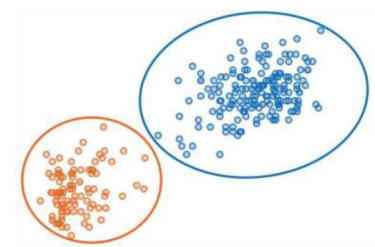
Zasada **podobieństwa** mówi, że ludzki mózg kategoryzuje obiekty na podstawie ich wspólnych atrybutów, takich jak kolor, kształt lub kierunek (np. Rysunek 9.2).



Rysunek 9.2. Podobieństwo jako zasada teorii Gestalt

Źródło: Schwabish (2021)

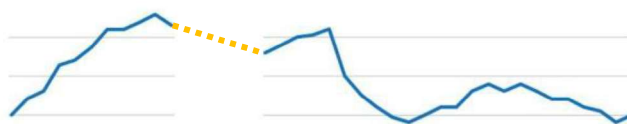
Zgodnie z zasadą **zamknięcia**, ograniczone obiekty są postrzegane jako grupa (np. Rysunek 9.3).



Rysunek 9.3. Zamknięcie jako zasada teorii Gestalt

Źródło: Schwabish (2021)

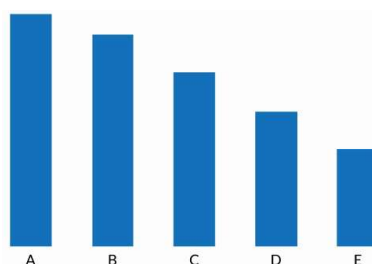
Zgodnie z zasadą **domknięcia**, nasz mózg ma tendencję do ignorowania luk i uzupełniania brakujących informacji w celu utworzenia kompletnej struktury. Analizując wykres liniowy, który zawiera brakujące dane, mamy tendencję do mentalnego wypełniania luk przy użyciu najprostszego podejścia (np. Rysunek 9.4).



Rysunek 9.4. Domknięcie jako zasada teorii Gestalt

Źródło: Schwabish (2021)

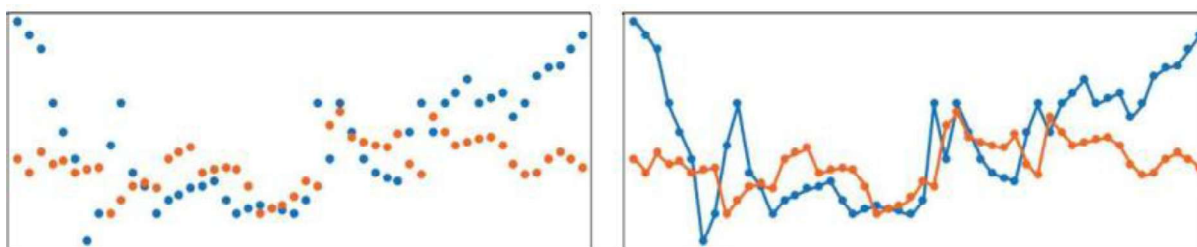
Zasada **ciągłości** sugeruje, że elementy ustawione w linii prostej lub gładkiej krzywej są postrzegane przez odbiorcę jako bardziej powiązane niż elementy, które nie leżą na linii lub krzywej (np. Rysunek 9.5).



Rysunek 9.5. Ciągłość jako zasada teorii Gestalt

Źródło: Schwabish (2021)

Opierając się na zasadzie **połączenia**, nasza percepcja kategoryzuje obiekty, które są ze sobą połączone, jako należące do tej samej grupy (np. Rysunek 9.6).



Rysunek 9.6. Połączenie jako zasada teorii Gestalt

Źródło: Schwabish (2021)

Istnieje bardzo ważna koncepcja wizualizacji danych, podzbiór teorii Gestalt zwany przetwarzaniem przeduwagowym. Schwabish (2021) wyjaśnia, że cechy przeduwagowe zwracają naszą uwagę na określony obszar wykresu lub obrazu.

Cechy te odnoszą się do jakości wizualnych, które ludzki system wizualny może postrzegać na wczesnych etapach przetwarzania wizualnego bez świadomej uwagi i które są zwykle mierzone w milisekundach. Jednym z atrybutów przeduwagowych używanych w tej książce jest kolor (niebieski) i waga (pogrubiony tekst). Kolejnym bardzo popularnym przykładem jest znalezienie określonej liczby w macierzy liczb (np. Wexler i in., 2017). Rysunek 9.7 przedstawia macierz liczb bez (po lewej stronie) i z (po prawej stronie) atrybutami przeduwagowymi.

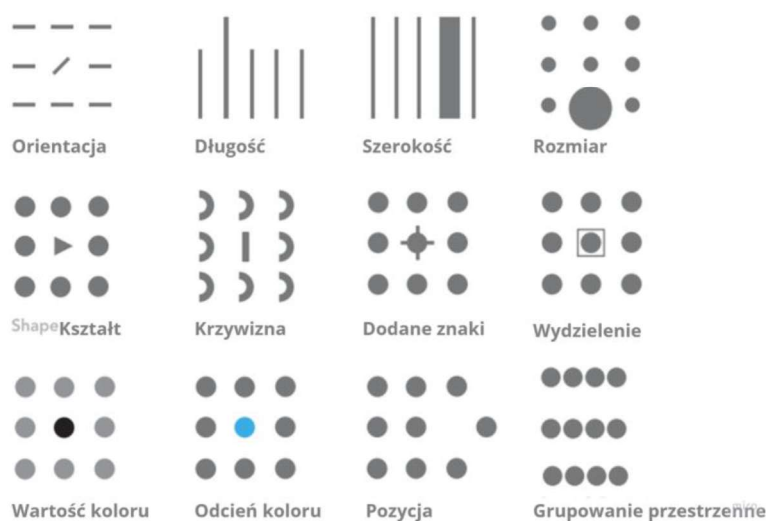
2	2	5	6	7	1	1	6	9	1
9	1	7	5	5	5	6	2	5	9
4	5	2	9	6	9	7	6	4	6
8	1	5	7	8	5	6	6	6	7
7	2	3	6	8	9	1	7	9	1
3	8	6	8	4	5	6	9	4	5
4	9	9	2	3	7	1	9	1	2
3	7	8	1	6	1	5	6	1	6
5	6	6	8	6	6	9	1	2	6
3	2	4	2	6	9	4	2	7	1

2	2	5	6	7	1	1	6	9	1
9	1	7	5	5	5	6	2	5	9
4	5	2	9	6	9	7	6	4	6
8	1	5	7	8	5	6	6	6	7
7	2	3	6	8	9	1	7	9	1
3	8	6	8	4	5	6	9	4	5
4	9	9	2	3	7	1	9	1	2
3	7	8	1	6	1	5	6	1	6
5	6	6	8	6	6	9	1	2	6
3	2	4	2	6	9	4	2	7	1

Rysunek 9.7. Wykorzystanie atrybutów przeduwagowych w wizualizacji danych

Źródło: Schwabish (2021)

Odgadnięcie, ile jest dziewiątek na lewej macierzy zajmuje dużo czasu. Ale tylko jedna zmiana w macierzy robi dużą różnicę. Zmieniony został tylko kolor - 9 są czerwone, a wszystkie inne liczby są jasnoszare. Kolor (w tym przypadku odcień) jest jednym z kilku atrybutów uwagi wstępnej. Rysunek 9.8 pokazuje przykłady niektórych atrybutów przeduwagowych często używanych w wizualizacji danych.



Rysunek 9.8. Rodzaje atrybutów przeduwagowych używanych w wizualizacji danych

Źródło: Schwabish (2021)

Atrakcyjne atrybuty umożliwiają widzom niemal natychmiastowe rozpoznawanie wzorców, wartości odstających lub ważnych punktów danych. Stosując strategie, które kierują wzrok odbiorcy na najważniejsze informacje, wizualizacje danych mogą znacznie poprawić komunikację i zrozumienie złożonych zestawów danych.

W następnym podrozdziale przeanalizujemy, w jaki sposób różne rodzaje danych i wnioski, które mają dostarczyć, wpływają na wybór metod wizualizacji. Od prostych wykresów słupkowych po bardziej złożone mapy cieplne lub wykresy punktowe, wybór odpowiedniej metody wizualizacji ma zasadnicze znaczenie dla zapewnienia, że dane nie tylko przyciągną uwagę, ale także skutecznie i dokładnie przekażą zamierzony komunikat.

9.3. Wybór właściwej metody wizualizacji

Pierwszym krokiem w wyborze odpowiedniej metody jest dokładne zrozumienie danych. Jakie są kluczowe wiadomości, które chcemy przekazać? Z jakim typem danych mamy do czynienia? Czy pracujemy z danymi szeregów czasowych, informacjami geograficznymi lub strukturami hierarchicznymi? Rodzaj wykorzystywanych danych może mieć znaczący wpływ na wybraną metodę wizualizacji. Na przykład mapy Choropleth najlepiej nadają się do wyświetlania danych geograficznych, podczas gdy wykresy liniowe mogą być bardziej odpowiednie dla danych szeregów czasowych.

Kontekst i oczekiwania odbiorców również odgrywają ważną rolę w wyborze metody wizualizacji. Publiczność techniczna może docenić szczegółowe i złożone wizualizacje, takie jak mapy cieplne lub diagramy sieciowe. Z drugiej strony, dla szerszej publiczności prostsze wykresy, takie jak wykresy słupkowe czy wykresy liniowe, mogą okazać się bardziej przystępne i angażujące.

Interaktywność to kolejny ważny aspekt, który należy wziąć pod uwagę. Interaktywne wizualizacje, takie jak dynamiczne dashboardy, umożliwiają użytkownikom eksplorowanie różnych poziomów danych poprzez filtrowanie, powiększanie i wybieranie określonych elementów. Ta interaktywność może prowadzić do głębszych spostrzeżeń, ponieważ użytkownicy mogą dostosować wizualizację do swoich konkretnych pytań.

Metodą wizualizacji nie zawsze muszą być wykresy. Może to być również tabela lub nawet zwykły tekst. Jak stwierdza Few (2012), celem tabel i wykresów jest skuteczne przekazywanie ważnych informacji i dostarczanie czytelnikowi ważnych, znaczących i przydatnych spostrzeżeń.

W kilku następnych podrozdziałach zostaną pokrótce wyjaśnione najpopularniejsze metody wizualizacji.

9.3.1. Prosty tekst

Nussbaumer Knaflic (2015) sugeruje użycie prostego tekstu, gdy do udostępnienia jest tylko jedna lub dwie liczby (Rysunek 9.9).

23%
wzrost sprzedaży w
porównaniu do 2022 r.

Rysunek 9.9. Prosty tekst w wizualizacji

Źródło: Schwabish (2021)

Według Schwabisha (2021) ten prosty tekst jest często określany jako BAN (ang. Big Ass Numbers). Są one zwykle używane do zwracania uwagi na kluczowe wskaźniki lub wskaźniki wydajności i zapewniają widzowi natychmiastowy dostęp do ważnych informacji. Podkreślając obszary wymagające uwagi lub działania, BAN zazwyczaj pomagają w podejmowaniu decyzji, pomagając użytkownikom skupić uwagę na ważnych aspektach danych (Tay, 2024). Chociaż BAN są proste, można je wzbogacić o subtelne elementy wizualne, takie jak kodowanie kolorami lub ikony, aby wskazać wydajność w stosunku do celów lub zmian w czasie. Na przykład czerwona strzałka w dół obok liczby sprzedaży może natychmiast wskazywać spadek, podczas gdy zielona strzałka w górę sygnalizuje wzrost.

9.3.2. Tabela

Tabele są istotną częścią wizualizacji danych, ponieważ zapewniają uporządkowany i przejrzysty sposób prezentacji danych liczbowych. Tabele są niezwykle przydatne, jeśli chodzi o prezentowanie szczegółowych informacji z precyzją i przejrzystością, mimo że mogą nie mieć takiego samego wpływu wizualnego jak wykresy lub mapy. Według Schwabisha (2020), w większości przypadków nie mają one na celu zapewnienia szybkiej wizualnej reprezentacji danych. Zamiast tego tabele są przydatne, gdy konieczne jest pokazanie dokładnych wartości danych lub szacunków. Chociaż nie są one idealną opcją do prezentacji dużej ilości danych lub na małej przestrzeni, dobrze zaprojektowana tabela może pomóc czytelnikowi znaleźć określone liczby, a także dostrzec trendy i wartości odstające. Few (2012) zauważa, że tabele są użyteczne dla odniesień i porównań jeden do jednego ze względu na ich prostą strukturę

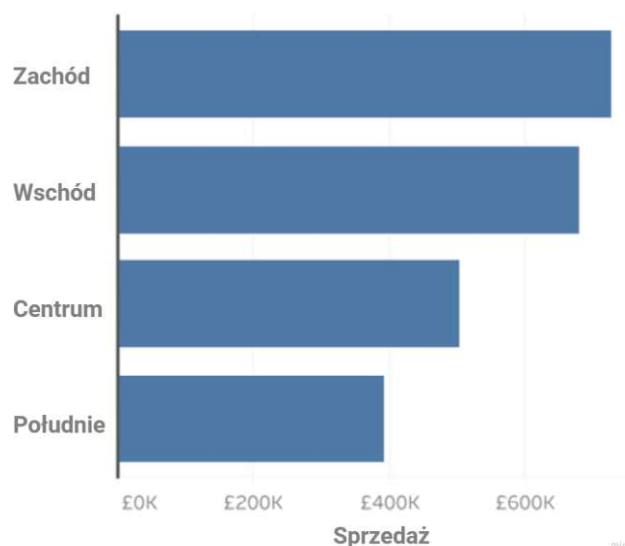
i fakt, że wartości ilościowe są wyrażone jako tekst, który możemy natychmiast zrozumieć bez konieczności tłumaczenia go.

Tabele powinny być sformatowane w następujący sposób (Schwabish, 2020; Nussbaumer Knaflitz, 2015):

- usuń wszystkie obramowania wokół tabeli;
- rozjaśnij linie siatki tak bardzo, jak to możliwe lub usuń je całkowicie;
- wyraźnie oddziel nagłówki od treści tabeli;
- wyrównaj tekst w tabeli i nagłówku do lewej strony, a liczb do prawej;
- używaj odpowiedniego poziomu szczegółowości danych (np. używaj liczb z jednym miejscem po przecinku, jeśli jest to wystarczające do zrozumienia danych).

9.3.3. Wykres słupkowy

Wykres słupkowy jest idealny do wyświetlania wartości liczbowych według grup lub kategorii (np. jeśli chcemy wyświetlić liczbę pracowników według działu). Może być wyrównany pionowo lub poziomo. Wyrównanie poziome (jak na Rysunku 9.10) jest zalecane, jeśli nazwy kategorii są zbyt długie lub jeśli istnieje zbyt wiele kategorii. Wykres słupkowy na rysunku 9.10. przedstawia sprzedaż (dane ilościowe) w podziale na regiony (dane jakościowe).



Rysunek 9.10. Wykres słupkowy w wizualizacji

Źródło: Wexler i in. (2021)

Według Few (2013) najskuteczniejszym sposobem przedstawienia miar związanych z dyskretnymi pozycjami na skali nominalnej lub porządkowej jest wykres słupkowy. Łatwo jest porównać poszczególne wartości, po prostu porównując wysokość słupków. Oś wykresu

słupkowego musi zaczynać się od zera. Jeśli oś zaczyna się od wartości innej niż zero, może to nadmiernie uwypuklić różnicę między słupkami i zniekształcić nasze postrzeganie wartości na wykresie słupkowym, które opiera się na długości słupków (Schwabish, 2021).

9.3.4. Wykresy liniowe

Wykres liniowy służy do przedstawienia zmiany wartości ilościowej, która leży na osi y, w czasie, który jest umieszczony na poziomej osi x. Yi (b.d.a) sugeruje, że wykres liniowy nie powinien zawierać więcej niż pięć linii. Ponadto nie jest konieczne uwzględnianie zerowej linii bazowej dla wykresu liniowego. Dopuszczalne jest rozszerzenie zakresu osi pionowej do punktu, w którym zmiany wartości są najbardziej pouczające, jeśli linia zerowa nie jest ani zrozumiała, ani pomocna.

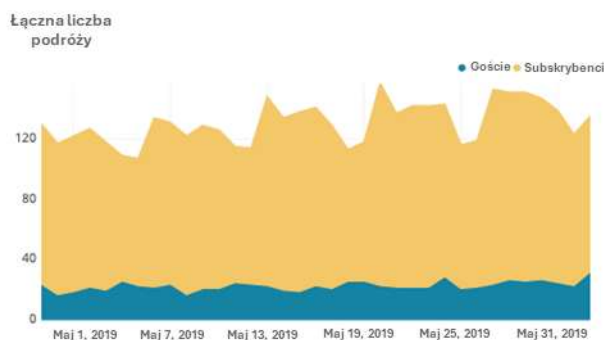


Rysunek 9.11. Wykres liniowy w wizualizacji

Źródło: Wexler i in. (2021)

Wykres liniowy na Rysunku 9.11 przedstawia sprzedaż (dane ilościowe) w okresie 4 lat i jest również podzielony na segmenty.

Wykres obszarowy (Rysunek 9.12), który jest wariantem wykresu liniowego, dodaje cień między linią a zerową linią bazową (Yi, b.d.a).



Rysunek 9.12. Wykres obszarowy w wizualizacji

Źródło: Yi (b.d.)

Wykres obszarowy można postrzegać jako skrzyżowanie wykresu liniowego i słupkowego, ponieważ wartości można interpretować nie tylko na podstawie ich pozycji w pionie, ale także na podstawie obszaru zacienionego między każdym punktem a linią bazową (Yi, n.d.).

9.3.5. Wykres punktowy

Wykres punktowy jest używany do sprawdzenia, czy istnieje związek między dwiema zmiennymi ilościowymi. Według The Data Visualisation Catalogue (b.d.) wzorce widoczne na wykresie punktowym można wykorzystać do interpretacji charakteru korelacji. Są to wartości: dodatnie (wartości rosną razem), ujemne (jedna wartość maleje, podczas gdy druga rośnie) lub zero (brak korelacji).



Rysunek 9.13. Wykres punktowy w wizualizacji

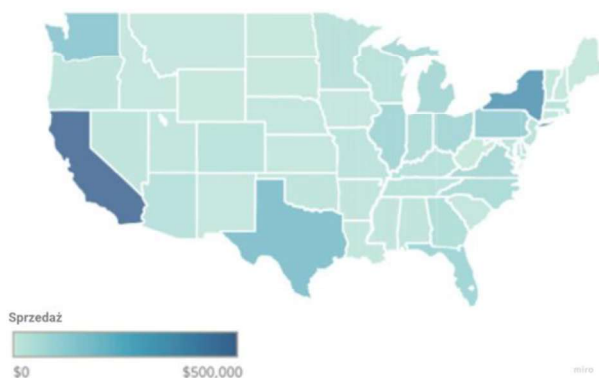
Źródło: Wexler i in. (2017)

Wykres punktowy na Rysunku 9.13 pokazuje związek między zyskiem a sprzedażą (obie zmienne ilościowe).

Według Yi (b.d.) bardzo ważne jest, aby wspomnieć, że na wykresie punktowym, nawet jeśli widzimy związek między dwiema zmiennymi, nie oznacza to, że zmiany jednej zmiennej powodują zmiany drugiej. Prowadzi to do powszechnie używanego w statystyce wyrażenia: „korelacja nie implikuje związku przyczynowego”.

9.3.6. Kartogram

Kartogram wykorzystuje różnice w cieniowaniu lub kolorowaniu w obrębie wcześniej zdefiniowanych obszarów, aby wskazać wartości lub kategorie w tych obszarach (Wexler i in., 2017). Według Schwabisha (2021) paleta kolorów na kartogramie jest łatwa do zrozumienia, mniejsze wartości odpowiadają jaśniejszym kolorom, a większe wartości ciemniejszym kolorom.



Rysunek 9.14. Kartogram w wizualizacji

Źródło: Wexler i in. (2017)

Kartogram na Rysunku 9.14 pokazuje sumę sprzedaży w różnych stanach USA.

9.3.7. Mapa cieplna

Mapa cieplna to wizualizacja danych w formacie tabelarycznym, gdzie kolorowe komórki reprezentują względną wielkość liczb (Nussbaumer Knafllic, 2015).

Ponieważ kolor jest kluczowym elementem tego typu wykresu, należy upewnić się, że wybrana paleta kolorów pasuje do danych. Najpopularniejszym rodzajem koloru jest kolor sekwencyjny, w którym ciemniejsze kolory korelują z wyższymi wartościami, a jaśniejsze kolory korelują z niższymi wartościami lub odwrotnie (Yi, b.d.b).

	Region A	Region B	Region C
Kategoria 1			
Kategoria 2			
Kategoria 3			

Kategoria 4			
Kategoria 5			

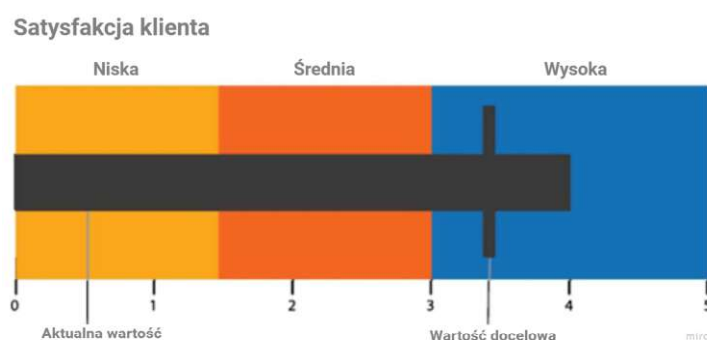
Rysunek 9.15. Mapa cieplna w wizualizacji

Źródło: Opracowanie własne na podstawie Nussbaumer Knaflic (2015)

Mapa cieplna na Rysunku 9.15 pokazuje różne wartości niektórych danych ilościowych (np. sprzedaży) według kategorii (w wierszach) i regionu (w kolumnach).

9.3.8. Wykres pociskowy

Wykres pociskowy został wynaleziony przez Stephena Few w 2005 roku (Few, 2013). Jest to w zasadzie wykres słupkowy z pojedynczym czarnym poziomym paskiem reprezentującym rzeczywistą wartość, dodatkowym (pionowym) znacznikiem wartości docelowej (którą chcemy osiągnąć) i zacienionymi obszarami w tle reprezentującymi skalę sukcesu (np. słaby, dobry, doskonały).



Rysunek 9.16. Wykres pociskowy w wizualizacji

Źródło: Schwabish (2021)

Wykres pociskowy na Rysunku 9.16 pokazuje, że chcemy osiągnąć ocenę zadowolenia klienta na poziomie 3,4 (w skali 5). Nasza obecna ocena zadowolenia wynosi 4, co jest wartością wyższą od wartości docelowej. W tle znajdują się trzy obszary zadowolenia klienta – niski (słaby), średni (dobry) i wysoki (doskonały).

Wybór odpowiedniej metody wizualizacji jest bardzo ważny dla skutecznej komunikacji, ale zasady projektowania, które kierują tymi technikami, również odgrywają kluczową rolę w przejrzystości i skuteczności prezentacji danych. W następnej sekcji przyjrzymy się znaczeniu układu, typografii, schematów kolorów i strategicznego wykorzystania przestrzeni, które mają kluczowe znaczenie dla uczynienia wizualizacji nie tylko estetycznymi, ale także łatwymi do zrozumienia i interpretacji.

9.4. Wytyczne dotyczące projektowania wizualizacji

Skuteczne projektowanie wizualizacji polega na zwiększeniu zdolności odbiorcy do zrozumienia i interakcji z danymi. Wiąże się to z utrzymaniem starannej równowagi między elementami estetycznymi a funkcjonalnością, przy czym wybór koloru, czcionki i układu odgrywa kluczową rolę w jasnym i skutecznym przekazywaniu informacji. Prostota powinna być również zasadą przewodnią. Jednak według Cairo (2013) grafika nie powinna upraszczać wiadomości. Powinna je wyjaśniać, podkreślać trendy, odkrywać wzorce i ujawniać realia, które wcześniej nie były widoczne.

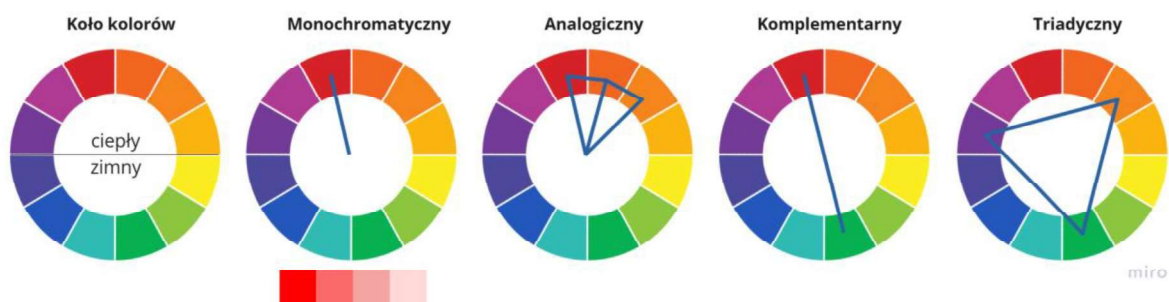
Częstą pułapką jest nadmierne komplikowanie wizualizacji zbyt wieloma elementami, które raczej dezorientują niż wyjaśniają. Celem jest uczynienie danych dostępnymi i zrozumiałymi dla docelowych odbiorców oraz zapewnienie, że wizualizacja służy swojemu celowi, którym jest informowanie i wspieranie podejmowania decyzji.

Przejrzystość i skuteczność wizualizacji danych można poprawić, usuwając niepotrzebne elementy rozpraszające uwagę. Nussbaumer Knaflic (2015) podkreśla, że każdy element, który nie dodaje wartości lub bezpośrednio nie wspiera zrozumienia danych, jest uważany za **bałagan**. Obejmuje to niepotrzebne linie siatki, nadmierne kolory, nieistotne punkty danych i zbyt złożone dekoracje wykresów. Elementy te mogą odwracać uwagę od kluczowych komunikatów, które dane mają przekazywać. Zaleca takie techniki, jak uproszczenie schematów kolorów, zminimalizowanie ilości tekstu i wykorzystanie białej przestrzeni. Dzięki strategicznemu wykorzystaniu białej przestrzeni można stworzyć wizualną hierarchię, która podkreśla kluczowe punkty danych i sprawia, że ogólna prezentacja jest bardziej przejrzysta i łatwiejsza do zrozumienia. Ponadto efektywne wykorzystanie białej przestrzeni może pomóc w stworzeniu zrównoważonego układu, który jest mniej zagracony i bardziej zorganizowany.

Kolor jest bardzo ważną częścią każdej wizualizacji danych. Służy nie tylko przyciąganiu uwagi, ale także organizowaniu informacji i skutecznemu przekazywaniu znaczenia. Prawidłowo użyty kolor może znacznie zwiększyć przejrzystość i wpływ wizualizacji. Istnieje kilka sugestii dotyczących skutecznego wykorzystania kolorów w wizualizacjach (Few, 2012; Cairo, 2013; Few, 2013; Nussbaumer Knaflic, 2015; Wexler i. in., 2017; Schwabish, 2021; Lidwell, 2023; Interaction Design Foundation, b.d.):

- **Wybierz odpowiednią kombinację kolorów:** korzystając z koła kolorów, można tworzyć wizualizacje, które są wizualnie zrównoważone i przyjemne dla oka. Istnieje kilka typowych kombinacji kolorów (Rysunek 9.17):
 - **Monochromatyczny** - jeden kolor w różnych odcieniach.
 - **Analogiczny** - trzy kolory znajdujące się obok siebie na kole kolorów. Kolory te są przyjemne dla oka i tworzą harmonijny projekt.
 - **Komplementarny** - dwa przeciwstawne kolory na kole kolorów. Kolory te są kontrastowe i powinny być używane do podkreślenia czegoś (np. wzrost - zielony/spadek - czerwony).
 - **Triadyczny** - trzy jednakowo odległe kolory na kole kolorów. Kolory te są dynamiczne i przyciągają uwagę.

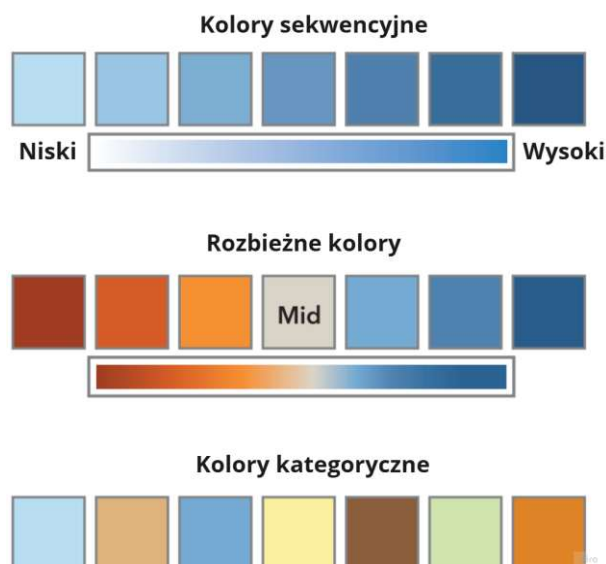
Ponadto, cieplejsze kolory powinny być używane dla elementów pierwszoplanowych, a chłodniejsze dla elementów tła.



Rysunek 9.17. Kombinacje kolorów

Źródło: Opracowanie własne na podstawie Lidwell (2023)

- **Wybierz odpowiednie schematy kolorów:** wybór schematu kolorów zależy od rodzaju danych, które mają być wizualizowane. **Sekwencyjny** schemat kolorów polega na użyciu tego samego koloru od jasnego do ciemnego odcienia. Jest on idealny do wyświetlania danych liczbowych, które przechodzą od niskiej do wysokiej wartości (np. sprzedaż według stanu). **Rozbieżny** schemat kolorów jest przydatny do podkreślania wartości powyżej lub poniżej punktu środkowego (np. zysku/straty). **Kategoryczny** schemat kolorów jest najlepszy dla danych kategorycznych, w których kolory muszą rozróżniać różne grupy bez sugerowania kolejności lub wartości (np. kategorie produktów). Rysunek 9.18 przedstawia różne schematy kolorów używane w wizualizacjach.



Rysunek 9.18. Schematy kolorów

Źródło: Wexler i in. (2017)

- **Używaj kolorów oszczędnie:** nadmierne użycie kolorów może powodować zamieszanie i utrudniać zrozumienie wykresu. Paleta kolorów powinna być ograniczona do tego, co ludzkie oko może szybko rozróżnić, czyli około pięciu różnych kolorów.
- **Rozważ ślepotę barw:** około 8% mężczyzn i 0,5% kobiet nie rozróżnia kolorów. Unikaj kombinacji kolorów, które są trudne do rozróżnienia dla użytkowników nierozróżniających kolorów, takich jak czerwony i zielony. Zamiast tych kolorów lepszym połączeniem jest pomarańczowy i niebieski.
- **Kolory powinny być spójne:** spójne użycie kolorów w wielu wizualizacjach pozwala odbiorcy łatwo zrozumieć i porównać dane. Po ustaleniu schematu kolorów dla określonych typów danych lub kategorii należy go zachować we wszystkich powiązanych wizualizacjach.
- **Użyj koloru, aby wyróżnić ważne dane:** kolor może być silnym wskaźnikiem tego, gdzie patrzeć. Użycie jasnego lub kontrastowego koloru może zwrócić uwagę na kluczowe punkty danych lub ustalenia, podczas gdy bardziej neutralne kolory mogą być używane do mniej krytycznych informacji. Niektórzy autorzy sugerują, że tworzenie jasnej, zrozumiałej wizualizacji powinno rozpocząć się od szarego koloru. Wszystkie elementy danych na wykresie (np. słupki na wykresie słupkowym lub linie na wykresie liniowym) powinny być szare. Następnie dodać należy etykiety i kolor tylko dla elementów, które chce się wyróżnić.
- **Zachowaj prostotę:** w dziedzinie wizualizacji danych zasada KISS, skrót od „Keep It Simple, Stupid”, jest bardzo istotna i pomocna. Odnosi się ona do korzystania

z prostych wykresów, ponieważ złożone wykresy lub zbyt szczegółowe wizualizacje mogą przytłoczyć użytkowników i utrudnić rozpoznanie kluczowych komunikatów lub punktów danych. Oznacza to również, że należy unikać wizualnego bałaganu i ograniczać niepotrzebne elementy wizualne, takie jak zbyt jasne kolory, czcionki i linie siatki. Stosując zasadę KISS, należy skupić się na samych danych, a nie na dekoracyjnych lub zbyt skomplikowanych elementach projektu.

Zasady te mają kluczowe znaczenie dla zapewnienia, że wizualizacje danych osiągną swój główny cel, jakim jest przekazywanie złożonych informacji w sposób dostępny i zrozumiały dla wszystkich odbiorców.

Z niniejszego rozdziału jasno wynika, że skuteczna wizualizacja danych jest kluczowym elementem w procesie podejmowania decyzji opartych na danych. W oparciu o zrozumienie kontekstu sytuacyjnego, w niniejszym rozdziale podkreślono znaczenie dostosowania wizualizacji do konkretnych potrzeb odbiorców docelowych. Poprzez szczegółową analizę różnych metod wizualizacji i zasad projektowania wykazano, że przemyślana wizualna reprezentacja danych jest ważna, aby umożliwić lepsze zrozumienie i komunikację złożonych informacji.

BIBLIOGRAFIA

1. Brush, K. (2022). Data visualization. TechTarget [dostępne: <https://www.techtarget.com/searchbusinessanalytics/definition/data-visualization>, dostęp 15.04.2024]
2. Cairo, A. (2013). The functional art: An introduction to information graphics and visualization. New Riders.
3. Data Visualisation Catalogue (b.d.). Scatterplot [dostępne: <https://datavizcatalogue.com/methods/scatterplot.html>, dostęp 17.04.2024]
4. Few, S. (2012). Show Me the Numbers. Analytics Press.
5. Few, S. (2013). Information dashboard design: Displaying data for at-a-glance monitoring. Analytics Press.
6. GeeksForGeeks (2024). What is Data Visualization and Why is It Important? dostępne: <https://www.geeksforgeeks.org/data-visualization-and-its-importance/>, dostęp 15.04.2024]
7. IBM (b.d.). What is data visualization? [dostępne: <https://www.ibm.com/topics/data-visualization>, dostęp 15.04.2024]
8. Interaction Design Foundation (b.d.). Color Theory [dostępne: <https://www.interaction-design.org/literature/topics/color-theory>, dostęp 18.04.2024]
9. Lidwell, W., Holden, K. & Butler, J. (2023). Universal Principles of Design, 3rd Edition. Quarto Publishing Group USA.
10. Nussbaumer Knaflic, C. (2015). Storytelling with data: A data visualization guide for business professionals. Wiley.
11. Schwabish (2020). Ten guidelines for better tables. Journal of Benefit-Cost Analysis, 11(2), pp. 151-178.
12. Schwabish (2021). Better data visualization: A guide for scholars, researchers and wonks. Columbia university press.
13. Tay, J. (2024). Effective use of BANs. Medium [dostępne: <https://medium.com/@e0373084/eye-catching-bans-88d29632e4fa>, dostęp 12.04.2024]
14. Wexler, S., Shaffer, J. & Cotgreave, A. (2017). The big book of dashboards: Visualizing your data using real-world business scenarios. Wiley.

15. Yi, M. (b.d.a). A complete guide to line charts. Atlassian [dostępne: <https://www.atlassian.com/data/charts/line-chart-complete-guide>, dostęp 17.04.2024]
16. Yi, M. (b.d.b). A complete guide to heatmaps. Atlassian [dostępne: <https://www.atlassian.com/data/charts/heatmap-complete-guide>, dostęp 17.04.2024]