**1 Idea *quantified self*: Zbieranie i analiza danych osobistych**

**1.1 Wprowadzenie i kluczowe zagadnienia**

Koncepcja *quantified self* odnosi się do zbierania i analizowania danych osobistych. Motywacje stojące za monitorowaniem mogą być różne – od poprawy jakości życia, poprzez podejmowanie bardziej świadomych decyzji, optymalizowanie zachowań i nawyków, eksplorowanie idei lub teorii, aż po czystą ciekawość. Ideę *quantified self* można podsumować frazą: „samopoznanie poprzez liczby”. Termin ten został po raz pierwszy wprowadzony przez Gary'ego Wolfa i Kevina Kelly'ego w 2007 roku, zainspirowanych obserwacjami ludzi, którzy śledzili ilościowe pomiary, takie jak waga, spożycie kalorii, wydatki czy nastrój[[1]](#footnote-1). Chociaż praktyka monitorowania danych osobistych jest stosowana od wielu lat jako narzędzie samodoskonalenia, to właśnie ostatnie osiągnięcia technologiczne – zwłaszcza w zakresie przechowywania i przetwarzania danych oraz zaawansowanych sensorów biometrycznych – wyniosły analitykę osobistą na nowy poziom popularności[[2]](#footnote-2).

Wielu badaczy, inżynierów i informatyków eksperymentowało z technologiami cyfrowymi, w szczególności komputerami ubieralnymi (en. *wearable computers*), w celu monitorowania osobistego. Jednym z najwcześniejszych pionierów był kanadyjski inżynier Steve Mann, często nazywany „ojcem komputerów ubieralnych”. Mann zdefiniował komputery ubieralne jako noszony system komputerowy, który jest zawsze włączony, gotowy i dostępny. W przeciwieństwie do tradycyjnych komputerów stacjonarnych, których głównym celem jest przetwarzanie danych, komputer ubieralny nie koncentruje się wyłącznie na obliczeniach. Jego założeniem jest wspomaganie użytkownika w wykonywaniu innych czynności, jednocześnie oferując funkcje komputerowe. W związku z tym komputer ubieralny ma na celu wspieranie intelektu użytkownika lub wzmacnianie jego zmysłów, zamiast wymagać pełnej uwagi i skupienia[[3]](#footnote-3).

Mann rozpoczął eksperymenty z komputerami noszonymi w latach 70. i zbudował swój pierwszy komputer ubieralny mając 12 lat. Do lat 80. używał tych urządzeń do rejestrowania informacji o swoich codziennych aktywnościach. Jego praca wykraczała poza proste monitorowanie; wyobrażał sobie technologię ubieralną jako sposób na stworzenie tego, co nazwał „rzeczywistością modyfikowaną” (en. „*mediated reality*”) – zdolność do zmiany lub ulepszania doświadczeń sensorycznych poprzez cyfrową augmentację[[4]](#footnote-4). W 1992 roku Mann założył Projekt Wearable Computing na MIT, gdzie kontynuował rozwój i doskonalenie swojej wizji technologii ubieralnej. W połowie lat 90. jego urządzenia były zdolne do ciągłego rejestrowania i transmitowania jego działań za pomocą czegoś, co nazwał „Wearable Wireless Webcam” – umożliwiając transmisję na żywo ze swojego życia. W 1998 roku Mann wynalazł smartwatcha, co stanowiło wczesną formę samomonitorowania, rejestrującą zarówno jego otoczenie, jak i dane osobiste, i skutecznie demonstrującą potencjał urządzeń ubieralnych w kontekście zbierania i analizy danych osobistych.

Historycznie, osoby zainteresowane monitorowaniem własnych danych korzystały z papieru i długopisu do śledzenia osobistych metryk. Jednakże rozwój technologii radykalnie zmienił sposób, w jaki zbieramy i analizujemy dane na swój temat. Nowoczesne smartfony są wyposażone w wiele sensorów, które umożliwiają zbieranie różnych zmiennych, takich jak liczba kroków, odległość czy lokalizacja. Rośnie liczba aplikacji mobilnych pomagających użytkownikom monitorować różne aspekty ich codziennego życia, w tym nastrój, spożycie jedzenia, aktywność fizyczną, dane finansowe i wiele innych.

Być może najważniejszym katalizatorem w rozwoju i ekspansji ruchu *quantified self* były innowacje w zakresie sensorów biometrycznych. Urządzenia ubieralne, takie jak smartwatche i opaski fitness, pozwalają teraz na monitorowanie wielu wskaźników zdrowotnych, takich jak zmienność tętna, poziom tlenu we krwi czy temperatura ciała. Urządzenia te wykorzystują algorytmy, które przetwarzają dane z sensorów zewnętrznych, takich jak akcelerometry i żyroskopy, przekształcając surowe dane ruchu w przydatne informacje, takie jak dzienna liczba kroków lub klasyfikacja wykonywanego ruchu (np. cios w boksie). Dodatkowo, wiele urządzeń ubieralnych jest wyposażonych w sensory monitorujące sen, oferując użytkownikom cenne informacje na temat jakości i długości ich odpoczynku.

Urządzenia te nie tylko zbierają surowe dane, ale także generują metryki końcowe, które agregują i upraszczają złożone dane dla użytkowników. Na przykład monitory snu mogą generować „oceny snu”, które podsumowują jego jakość, podczas gdy oceny gotowości analizują stan fizyczny użytkownika i przygotowanie na nadchodzący dzień na podstawie danych z sensorów biometrycznych. Wskaźniki te dostarczają wartościowych, użytecznych informacji, przyczyniając się do rosnącej popularności ruchu *quantified self* i umożliwiając ludziom podejmowanie bardziej świadomych decyzji dotyczących zdrowia i stylu życia.

**1.2 Powody i motywacje do zbierania i analizy danych osobistych**

Wiele badań poświęcono analizie demografii oraz motywacji osób angażujących się w zbieranie i analizę danych osobistych. Na przykład w raporcie *Connected Life Report* z 2014 roku, przygotowanym przez Nielsen, wskazano, że młodzi dorośli w wieku od 25 do 34 lat najczęściej korzystają z opasek fitness lub aplikacji mobilnych do śledzenia danych, stanowiąc 40% próby badawczej. Raport ujawnił również, że kobiety w wieku od 30 do 39 lat dominowały wśród użytkowników aplikacji zdrowotnych i fitness. Dodatkowo zauważono związek między posiadaniem urządzeń ubieralnych a wyższym poziomem dochodów — co trzeci użytkownik opasek fitness deklarował dochód gospodarstwa domowego na poziomie co najmniej 100 000 USD rocznie[[5]](#footnote-5).

W innym badaniu, naukowcy z Uniwersytetu Waszyngtońskiego i Microsoftu przyjrzeli się praktykom entuzjastów self-trackingu, których nazywają *Quantified-Selfers*, w artykule pt. *Understanding Quantified-Selfers’ Practices in Collecting and Exploring Personal Data*. Analizowali 52 nagrania ze spotkań społeczności *quantified self*, zbierając dane jakościowe i ilościowe od tzw. „użytkowników ekstremalnych” — osób, które wykazywały wysoką motywację pomimo wielu wyzwań, często tworząc własne rozwiązania alternatywne. Zdaniem autorów, użytkownicy ci dostarczali cennych informacji w szerszym kontekście ruchu *quantified self*. Badanie wykazało, że 79% uczestników stanowili mężczyźni, jednak autorzy podkreślili, że w ogólnej populacji self-tracking jest równomiernie rozłożone między płcie, co jest zgodne z wynikami raportu *Connected Life Report*. Pod względem zawodowym 40% uczestników pracowało w startupach, 37% zadeklarowało zawód inżyniera oprogramowania, a inni byli analitykami danych lub inżynierami elektrycznymi. Najczęściej monitorowanymi zmiennymi były aktywność fizyczna, spożycie jedzenia, waga, sen i nastrój[[6]](#footnote-6). Motywacje do samodzielnego zbierania i analizy danych można podzielić na cztery główne kategorie:

* Poprawa zdrowia: zarządzanie lub leczenie stanów chorobowych, ale również osiągniecie celów, takich jak utrata wagi.
* Ulepszanie życia: zwiększenie wydajności pracy, nauki lub związków międzyludzkich.
* Eksploracja: próbowanie nowych rzeczy i czerpanie przyjemności z procesu.

Badacze wskazali również trzy powszechne błędy popełniane przez osoby zbierające i analizujące dane osobiste:

1. Zbieranie zbyt wielu zmiennych: może prowadzić do zmęczenia lub zniechęcenia, zwłaszcza gdy zarządza się wieloma źródłami danych.
2. Brak kontekstowego zbierania danych: skupienie się wyłącznie na objawach bez uwzględniania kontekstu lub czynników środowiskowych ogranicza przydatność danych.
3. Brak naukowego podejścia: niewystarczająca dbałość o metodologię może prowadzić do mylących wniosków.

Wnioski te podkreślają zarówno potencjał, jak i wyzwania związane z monitorowaniem danych osobistych, ilustrując zdolność tej praktyki do poprawy jakości życia, a jednocześnie wskazując obszary wymagające udoskonalenia.

**1.3 Technologie i urządzenia ubieralne**

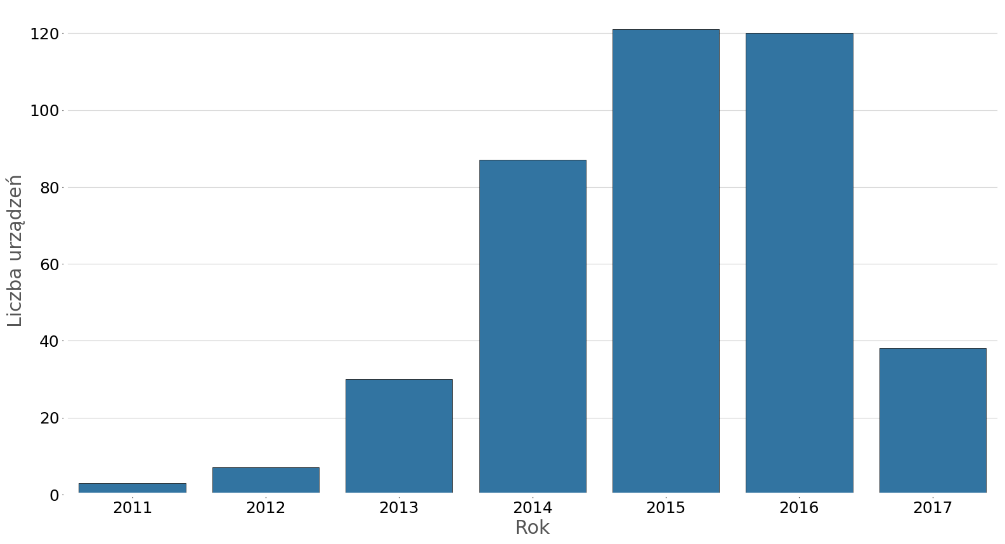
**1.3.1 Sensory w urządzeniach ubieralnych**

Sensory wykorzystywane w urządzeniach ubieralnych można podzielić na wewnętrzne i zewnętrzne w zależności od rodzaju zbieranych danych. Sensory zewnętrzne są umieszczone na powierzchni ciała, ale nie wymagają bezpośredniego kontaktu ze skórą i mierzą zjawiska środowiskowe, które zachodzą poza ciałem użytkownika. Ich główną zaletą jest łatwość w zakładaniu i stosowaniu, ponieważ nie wymagają bezpośredniego kontaktu z ciałem. Dodatkowo, sensory te zapewniają wygodę użytkowania, ponieważ często są niewidoczne lub minimalizują ingerencję w codzienne czynności. Przykładem takich sensorów mogą być akcelerometry, żyroskopy czy czujniki GPS, które są wykorzystywane do monitorowania ruchu, lokalizacji i orientacji użytkownika. Jednak sensory zewnętrzne mogą napotkać pewne ograniczenia, takie jak utrata zasięgu GPS w zamkniętych pomieszczeniach, jak np. w tunelach, co może wpłynąć na wiarygodność pomiarów.

Z kolei sensory wewnętrzne wymagają bezpośredniego kontaktu ze skórą użytkownika i zbierają dane dotyczące zjawisk zachodzących wewnątrz ciała. Przykładami takich sensorów są czujniki mierzące tętno, saturację krwi,nasycenie tlenem mięśni szkieletowych czy częstotliwość oddechów. Sensory wewnętrzne umożliwiają precyzyjniejsze monitorowanie parametrów, które mają bezpośredni wpływ na stan zdrowia i wydolność organizmu. Przykładowo, monitorowanie tętna może pełnić funkcję wskaźnika relatywnej intensywności wysiłku fizycznego, a także być wykorzystywane do oceny poziomu stresu, co daje cenną informację o kondycji psychofizycznej użytkownika.

Nowoczesne urządzenia ubieralne są najczęściej wyposażone zarówno w sensory wewnętrzne, jak i zewnętrzne. W 2017 roku naukowcy z Uniwersytetu w Tromsø przeprowadzili badanie, w którym zidentyfikowali urządzenia ubieralne i zgromadzili dane opisujące urządzenia ubieralne wypuszczone na rynek w okresie od 2011 do połowy 2017. W ramach badania zebrano dwanaście zmiennych, w tym między innymi nazwę urządzenia, rok wydania oraz obsługiwane sensory. Naukowcy łącznie zgromadzili dane dotyczące 423 urządzeń, jednocześnie informując o prawdopodobnej niekompletności zestawu danych, wynikającej ze względów praktycznych[[7]](#footnote-7).

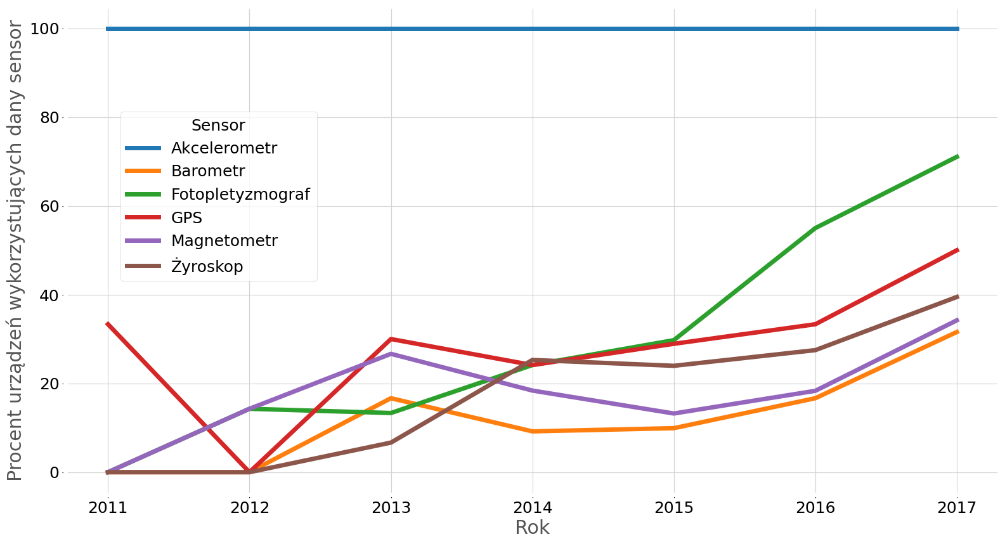
Na wykresie 1 przedstawiono liczbę urządzeń uwzględnionych w każdym roku.

*Wykres 1. Liczba urządzeń ubieralnych wprowadzanych na rynek w latach 2011-2017(I połowa)*

*Źródło: Opracowanie własne na podstawie André Henriksen i in., „Dataset of Fitness Trackers and Smartwatches to Measuring Physical Activity in Research”.*

Można zaobserwować wyraźny wzrost liczby wprowadzanych na rynek urządzeń, z wyjątkiem okresu 2016-2017, co może być wynikiem ograniczenia zbierania danych do połowy 2017 roku. Wykres ten ilustruje rosnącą popularność urządzeń ubieralnych oraz rozwój kultury *quantified self*.

Wykres 2 przedstawia procent urządzeń ubieralnych, które wykorzystywały dany sensor w danym roku.

*Wykres 2. Procent urządzeń ubieralnych wykorzystujących poszczególne sensory w latach 2011-2017*

*Źródło: Opracowanie własne na podstawie André Henriksen i in., „Dataset of Fitness Trackers and Smartwatches to Measuring Physical Activity in Research”.*

Generalna tendencja dla wszystkich sensorów wykazuje wzrost w badanym okresie, co świadczy o rosnącej funkcjonalności urządzeń ubieralnych, które stają się coraz bardziej uniwersalne i precyzyjne. Akcelerometr pozostawał niezmiennie obecny we wszystkich urządzeniach w latach 2011-2017 (Dane były zbierane tylko dla urządzeń bazujących na akcelerometrze). Niemniej jednak akcelerometr cechuje się niską ceną oraz doskonałymi właściwościami predykcyjnymi. Warto zwrócić szczególną uwagę na wzrastającą liczbę urządzeń wykorzystujących fotopletyzmograf, sensor wewnętrzny, który jest używany do estymacji tętna oraz zmienności rytmu serca. Wzrost ten może sugerować rosnące zainteresowanie monitorowaniem parametrów fizjologicznych użytkowników, co może być związane z rosnącą popularnością monitorowania zdrowia i kondycji.

**1.3.2 Aplikacje urządzeń ubieralnych**

Być może największym zastosowaniem technologii ubieralnych jest sport. Miliony amatorskich sportowców korzystają z tych urządzeń, aby oceniać swoje postępy. Jeszcze bardziej interesujące jest jednak wykorzystanie urządzeń ubieralnych w profesjonalnym sporcie. Organizacje, którym zależy na uzyskaniu nawet najmniejszej przewagi nad konkurencją, mogą wykorzystywać technologie ubieralne do zarządzania stresem, monitorowania obciążenia i zmęczenia zawodników oraz zapobiegania kontuzjom. Istnieją również urządzenia projektowane specjalnie z myślą o konkretnych dyscyplinach sportowych. Przykładem takiego instrumentu jest VERT — urządzenie monitorujące wertykalne skoki zawodników, najczęściej siatkarzy. VERT wykorzystuje zewnętrzny sensor, konkretnie akcelerometr, do generowania danych dla pojedynczej sesji takich jak: wysokość najwyższego skoku, liczba skoków, czas aktywności w minutach oraz liczba skoków powyżej 38 centymetrów.

Akcelerometr to urządzenie służące do pomiaru przyspieszenia, czyli zmiany prędkości obiektu w czasie. Akcelerometr może mierzyć przyspieszenie w trzech osiach: X, Y, oraz Z, co pozwala na precyzyjne śledzenie ruchu ciała w przestrzeni trójwymiarowej.

* Oś X: Reprezentuje ruch w poziomie (w lewo/w prawo).
* Oś Y: Reprezentuje ruch w pionie (w górę/w dół).
* Oś Z: Reprezentuje ruch w poziomie (do przodu/do tyłu).

Niektóre urządzenia ubieralne używane w sporcie konwertują dane z akcelerometru na metrykę nazywaną „obciążeniem zawodnika” (en*. „player load”*). Obciążenie zawodnika może być wykorzystywane do oceny wysiłku podczas treningu oraz do zarządzania nakładem treningowym w trakcie sezonu w celu uniknięcia przetrenowania lub zbyt małego obciążenia treningowego, a także w celu zapobiegania kontuzjom. Przykładem urządzenia obliczającego obciążenie zawodnika jest Catapult, wykorzystywane w sportach drużynowych, takich jak piłka nożna, rugby czy hokej. Skumulowane obciążenie zawodnika w okresie od 0 do n uwzględnia wszystkie trzy osie odczytów akcelerometru i stanowi sumę natychmiastowych obciążeń zawodnika[[8]](#footnote-8)

gdzie:

* t to kolejne momenty czasu w trakcie pomiaru (od t=0 do t=n).

Być może najważniejszą aplikacją urządzeń ubieralnych w sporcie jest zapobieganie kontuzjom. Jednym ze sposobów rozwiązania tego problemu jest monitorowanie proporcji obciążenia zawodnika w krótkim okresie do obciążenia w długim okresie. Takie obliczenie jest znane jako proporcja obciążenia gwałtownego do przewlekłego (en. *acute to chronic workload ratio, ACWR*). Obciążenie gwałtowne często odnosi się do nakładu treningowego w ostatnim tygodniu, natomiast obciążenie przewlekłe często do nakładu z ostatnich 3–6 tygodni[[9]](#footnote-9).

gdzie:

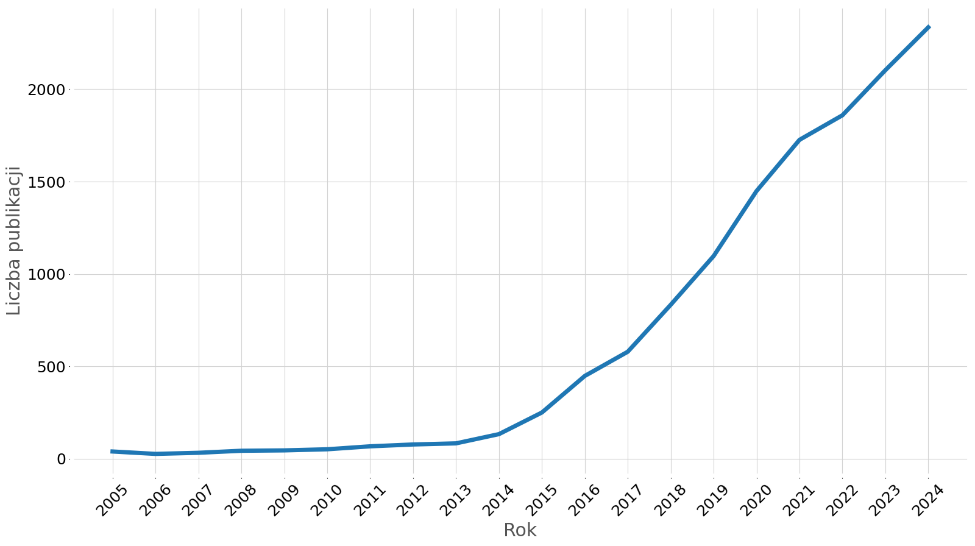
Sytuacje, w których ACWR przekracza wartość 1.5, są powszechnie uznawane za okoliczności wiążące się ze zwiększonym ryzykiem kontuzji, ponieważ gwałtowne zmiany w obciążeniu treningowym mogą prowadzić do przeciążenia tkanek i w konsekwencji urazów. Z kolei ACWR w przedziale od 0.80 do 1.30 bywa traktowane jako obszar, w którym ryzyko kontuzji jest zredukowane, sugerując, że organizm ma wystarczającą zdolność do adaptacji do obciążeń w długim okresie, przy jednoczesnym utrzymaniu odpowiedniej intensywności treningu w krótkim okresie[[10]](#footnote-10).

Pomimo popularności i szerokiego zastosowania ACWR, istnieją powody do sceptycznego podejścia do tego wskaźnika jako predyktora lub przyczyny kontuzji wśród sportowców. ACWR jest uproszczonym modelem, który nie uwzględnia indywidualnych zdolności adaptacyjnych zawodników ani złożonych interakcji między zmiennymi, takimi jak poziom zmęczenia, stan fizyczny, psychologiczne predyspozycje czy inne czynniki zewnętrzne[[11]](#footnote-11).

Urządzenia i technologie ubieralne okazują się być przydatne nie tylko w celu optymalizacji treningu, regeneracji lub zachowań sportowców. Wykazują również potencjał w usprawnieniu i automatyzacji metod sędziowania. Przykładowo, naukowcy z Uniwersytetu Griffitha zaproponowali użycie sensorów zewnętrznych, akcelerometrów i żyroskopów, w celu automatycznej klasyfikacji pięciu typów ciosów w boksie. Używając sześciu różnych typów modeli klasyfikacji nadzorowanej, osiągnęli średnią dokładność klasyfikacji wynoszącą 90% +- 12%. Autorzy sugerują, że takie wykorzystanie technologii mogłoby wspierać sędziów podczas zawodów, umożliwiając im bardziej precyzyjną ocenę wyników walki oraz zwiększając obiektywność decyzji, tym samym wpisując się w szerszy trend integracji systemów automatycznej analizy danych, mającymi na celu zwiększenie transparentności i obiektywizmu w sporcie[[12]](#footnote-12).

Kolejną potencjalną aplikacją technologii ubieralnych w sporcie jest automatyczna klasyfikacja ćwiczeń oraz liczenie powtórzeń. W badaniu przeprowadzonym przez studenta Wolnego Uniwersytetu w Amsterdamie, pięciu uczestników wykonywało ćwiczenia ze sztangą (wyciskanie sztangi leżąc, przysiad, wyciskanie sztangi nad głową, wiosłowanie ze sztangą, martwy ciąg). Uczestnicy mieli zamocowane urządzenia na nadgarstkach, które zbierały dane z akcelerometrów i żyroskopów podczas wykonywania ćwiczeń ze sztangą. Autor badania testował różne modele klasyfikacji ćwiczeń. Ostatecznie najlepszy model osiągnął dokładność 98.51% na zestawie danych testowych, czyli na obserwacjach, które nie były wykorzystywane podczas procesu trenowania. Zdaniem autora, automatyczna klasyfikacja wykonywanych ćwiczeń wykazuje duży potencjał komercyjny, umożliwiając integrację z urządzeniami i aplikacjami obsługującymi technologie ubieralne oraz tworzenie cyfrowych trenerów personalnych[[13]](#footnote-13).

Kolejnym zastosowaniem urządzeń ubieralnych jest medycyna i służba zdrowia. Urządzenia te pozwalają na ciągłe monitorowanie oznak fizjologicznych pacjentów i stają się coraz dokładniejsze w pomiarach. Globalne trendy oraz liczba publikacji naukowych wskazują na gwałtowny rozwój wykorzystania urządzeń ubieralnych w medycynie, sugerując, że w przyszłości staną się one istotnym elementem systemu opieki zdrowotnej[[14]](#footnote-14).

*Wykres 5. Liczba publikacji na temat urządzeń ubieralnych w kontekście opieki zdrowotnej i medycyny na PubMed (2005–2024)*

*Źródło: Opracowanie własne na podstawie danych z PubMed[[15]](#footnote-15)*

Przykładem urządzeń używanych w służbie zdrowia są Google Glass, które wspierają proces operacji w czasie rzeczywistym. Robotyka ubieralna, zwłaszcza urządzenia wspomagające górne i dolne kończyny, zwiększają efektywność procesów rehabilitacyjnych, w szczególności u pacjentów po udarze. Internet rzeczy (IoT) łączy urządzenia medyczne, umożliwiając bardziej płynne monitorowanie i opiekę poza tradycyjnymi miejscami, takimi jak szpitale. Dodatkowo, czujniki ubieralne i inteligentne tekstylia przyczyniają się do dalszego postępu w medycynie, umożliwiając ciągłe monitorowanie stanu pacjenta i wspieranie go, szczególnie w przypadku chorób takich jak choroba Parkinsona. Pacjenci są w stanie na bieżąco monitorować wiele aspektów swojego życia, co pozwala na łatwiejsze dzielenie się tymi informacjami z lekarzami[[16]](#footnote-16).

PGHD (en. *Patient-Generated Health Data*) to termin odnoszący się do danych generowanych przez pacjentów. Wzrost popularności samomonitorowania sprzyja powstawaniu „cyfrowej kultury pacjentów”, w której pacjenci odgrywają bardziej aktywną rolę w zarządzaniu swoim zdrowiem. PGHD umożliwia pacjentom większe zaangażowanie i lepsze zrozumienie swojego stanu zdrowia, szczególnie w okresach między wizytami u lekarzy. PGHD pozwala także przekazywać dostawcom opieki zdrowotnej ciągłe dane, które dostarczają cennych informacji o stylu życia oraz czynnikach psychospołecznych, takich jak aktywność fizyczna, sen i dieta. Połączenie danych generowanych przez pacjentów z tradycyjnymi danymi medycznymi daje szansę na bardziej kompleksową i spersonalizowaną opiekę zdrowotną, co może przyczynić się do poprawy wyników leczenia oraz redukcji kosztów systemów ochrony zdrowia. PGHD ma potencjał, aby zrewolucjonizować tradycyjne modele opieki zdrowotnej, dając pacjentom większą kontrolę nad ich zdrowiem i procesem leczenia[[17]](#footnote-17).

**2 Metodologia**

**2.1 Eksploracyjna analiza danych (EDA)**

Eksploracyjna analiza danych (*en. Exploratory data analysis, EDA*)) to podejście do analizy danych zakładające istnienie gotowego zestawu danych wtórnych, który może zostać wykorzystany w projektach naukowych, biznesowych lub personalnych. Celem EDA jest stopniowe pogłębianie zrozumienia zbioru danych poprzez zastosowanie odpowiednich metod analitycznych. Stanowi ona kluczowy, początkowy etap procesu analizy danych. W ramach eksploracyjnej analizy danych stosuje się m.in. wizualizację danych, czyszczenie zbioru (np. identyfikację wartości brakujących, zduplikowanych, skrajnych lub nieprawidłowych), generowanie statystyk opisowych dla zmiennych oraz transformacje danych w celu ich lepszego dostosowania do późniejszych metod modelowania statystycznego. EDA nie tylko dostarcza informacji na temat kolejnych kroków w analizie, lecz także pozwala na identyfikację możliwości oraz ograniczeń związanych z danym zbiorem. Do eksploracyjnej analizy danych można również zaliczyć grupowania obserwacji, takie jak algorytm K-średnich, który klasyfikuje podobne do siebie jednostki w odrębne grupy.

Wizualizacja danych stanowi podstawę eksploracyjnej analizy danych. Pozwala na identyfikowanie kształtów rozkładów zmiennych oraz wstępne zrozumienie związków pomiędzy zmiennymi. Wizualizacja danych jest istotna, ponieważ wykorzystuje ludzką zdolność do identyfikacji skomplikowanych wzorców w formach graficznych[[18]](#footnote-18). Istotnym zadaniem jest odpowiedni dobór typu wykresu do zmiennej, lub zmiennych które są poddawane wizualizacji. Wykresy powinny być czytelne, zrozumiałe ale jednocześnie przenoszące jak najwięcej informacji. Równie ważny jest dobór wizualizacji do grupy docelowych odbiorców.

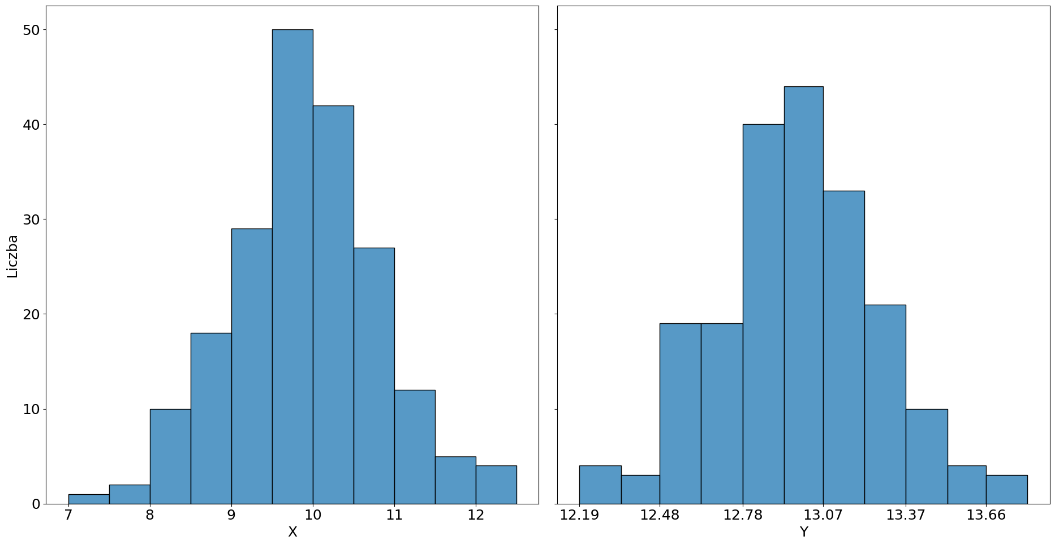
Czyszczenie danych jest kolejnym podstawowym zadaniem EDA. Większość zestawów danych generowanych współcześnie zawiera brakujące wartości, błędy, wartości skrajne. Na szczęście w dobie komputerów o wysokich mocach obliczeniowych, sama identyfikacja wartości brakujących nie jest problemem (ciężej jest definitywnie zidentyfikować błędy lub wartości skrajne). Bardziej istotna jest decyzja o traktowaniu tych wartości. Jeżeli jest ich na tyle mało, że usunięcie wierszy zawierających braki nie spowoduje dużej straty informacji, to taka decyzja może mieć sens. Istnieją jednak, lepsze, choć nieperfekcyjne metody na imputację wartości brakujących.

Najprostszą metodą jest imputacja brakującej wartości odpowiednią dla danej zmiennej miarą tendencji centralnej, czyli przykładowo średnią arytmetyczną dla cechy ilościowej ciągłej lub dominantą dla cechy jakościowej opisującej kategorię lub klasy. Łatwość i szybkość stosowania tej metody sprawia, że jest ona częstym wyborem, pomimo oczywistych ograniczeń. Próbą polepszenia imputacji miarą tendencji centralnej może być zawężanie kryterium obliczania wartości. Przykładowo użycie innej zmiennej jakościowej w celu podziału zbioru danych na grupy i na tej podstawie liczenie osobnej miary dla każdej grupy. Bardziej zaawansowaną, lecz potencjalnie skuteczniejszą metodą imputacji brakujących danych jest wykorzystanie modeli statystycznych. Jest to szczególnie użyteczne w sytuacjach, gdy brakujące wartości występują w jednej kolumnie. W takim przypadku można traktować tę kolumnę jako zmienną zależną (Y) i opracować model, który estymuje brakujące wartości na podstawie pozostałych dostępnych danych[[19]](#footnote-19).

Wartości skrajne to obserwacje znacząco odbiegające od typowego wzorca danych lub niepowiązane z innymi obiektami w zbiorze. Ich obecność może w niektórych przypadkach nadmiernie wpływać na analizę, prowadzić do błędnych wniosków i ostatecznie nietrafionych decyzji[[20]](#footnote-20). Pomimo tego, że identyfikacja wartości skrajnych jest bardziej zaawansowanym zadaniem od znajdowania brakujących wartości, istnieje wiele metod na ich identyfikację.

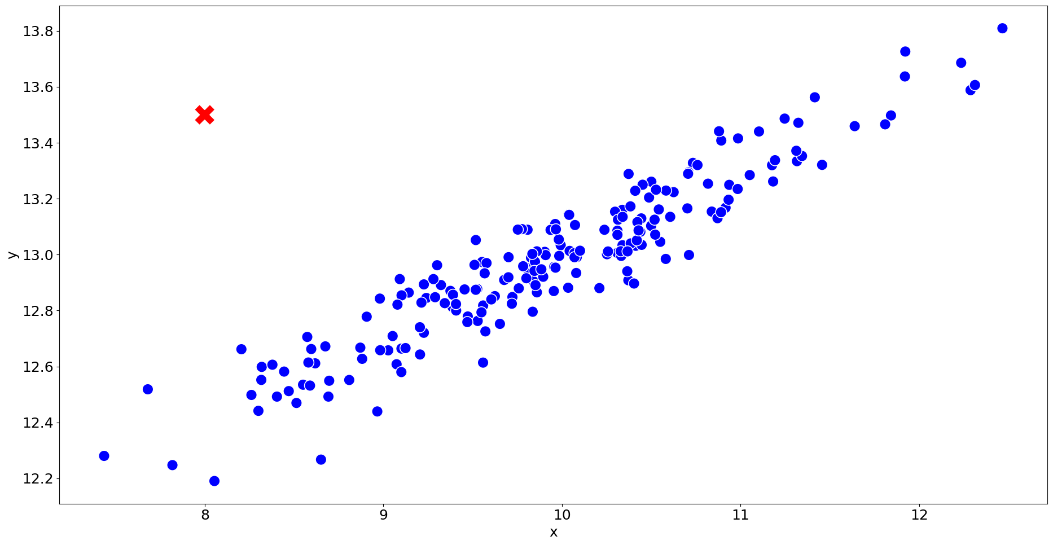
Jedną z najprostszych sformalizowanych metod identyfikacji wartości skrajnych jest metoda rozstępu międzykwartylowego. Wartości skrajne oznaczane na wykresach pudełkowych są rozpoznawane przy pomocy tej metody. Obserwacja zostanie uznana za skrajną, jeżeli wartość badanej cechy znajduje się w odległości 1,5 razy rozstęp międzykwartylowy (IQR) poniżej pierwszego kwartylu (Q1) lub powyżej trzeciego kwartylu (Q3). Matematycznie granice wykrywania wartości skrajnych można wyrazić jako[[21]](#footnote-21):

Wizualizacja danych może również dostarczać informacji w kontekście identyfikacji wartości skrajnych. Zaobserwowanie stosunkowo mało licznych przedziałów w dalekich ogonach rozkładu zmiennej wizualizowanej przy pomocy histogramu może sygnalizować obecność wartości skrajnych. Wykresy pudełkowe dają szybki obraz identyfikowanych wartości skrajnych. Wykresy rozrzutu pomagają w identyfikacji wartości skrajnych w kontekście dwóch zmiennych. Możliwe jest zobrazowanie potencjału podejścia wizualizacji danych w celu identyfikacji wartości skrajnych w dwóch wymiarach, przy użyciu dwóch wygenerowanych syntetycznych zmiennych o rozkładzie normalnym, które są ze sobą silnie skorelowane.

*Wykres 6. Rozkład syntetycznych zmiennych x i y.*

*Źródło: Opracowanie własne*

Zmienne oglądane w izolacji nie wykazują znaków występowania wartości skrajnych. Dopiero po obejrzeniu dwóch zmiennych jednocześnie łatwo można zidentyfikować nietypową obserwację, która istotnie odbiega od trendu, pomimo tego, że x=8 i y=13,5 nie są wartościami nietypowymi w kontekście zmiennych badanych pojedynczo.

*Wykres 7. Związek zmiennych x i y.*

*Źródło: Opracowanie własne*

Przykład ilustruje potencjał wizualizacji danych w kontekście identyfikacji wartości skrajnych, jednocześnie pokazując jej ograniczenia. Możliwości wizualizacji w praktyce kończą się na dwóch lub trzech wymiarach, dlatego metody analizy wielowymiarowej dają w wielu przypadkach skuteczniejsze oraz bardziej formalne wyniki identyfikacji wartości skrajnych.

Istotnym jest zdeterminowanie czy zidentyfikowana wartość skrajna jest wynikiem błędu (np. błędnego wprowadzenia przez człowieka lub wadliwego odczytu sensora), czy też stanowi rzeczywistą wartość, opisującą obserwację znacząco odbiegającą od pozostałych w zbiorze danych. Zdrowy rozsądek oraz wiedza ekspertów danej dziedziny pomaga w identyfikacji wartości skrajnych powstałych w wyniku błędu. Prostym przykładem może być zbiór danych dotyczący wzrostu ludzi, w którym znajduje się obserwacja o wartości -150 centymetrów. Wzrost jest zmienną ograniczoną do wartości nieujemnych, więc logicznym wnioskiem jest uznanie tej obserwacji za błędną. W przypadku zidentyfikowania wartości błędnej rozsądnym jest usunięcie całego wiersza obserwacji, lub usunięcie wartości jedynie błędnej zmiennej i użycie odpowiedniej metody imputacji. Z kolei w sytuacji, gdy wartość skrajna nie jest wynikiem błędu, lecz rzeczywistą cechą danej obserwacji, jej usunięcie może prowadzić do utraty istotnej informacji. Atrakcyjnością usunięcia takiej wartości jest potencjalne lepsze przygotowanie danych do danego modelu, który został wybrany jako następny krok analizy. Jednakże, celem eksploracji danych jest między innymi właśnie poprawienie świadomości dotyczącej charakterystyki danego zbioru danych oraz wybór następnych kroków w analizie. Dlatego zamiast ignorować wartości skrajne, bardziej zasadnym podejściem jest transparentne uwzględnienie ich w analizie oraz jawne zakodowanie założenia występowania wartości skrajnych poprzez odpowiedni wybór narzędzi statystycznych, które uwzględniają ich wpływ na estymację i niepewność predykcji.

Jeżeli wcześniejsza eksploracja danych wykazała, że w zbiorze mogą występować podgrupy, to przydatne mogą się okazać metody nienadzorowanej klasyfikacji, których celem jest przypisanie każdej obserwacji do grupy, w której znajdują się obserwacje jak najbardziej podobne do siebie. Każda grupa powinna być jak najbardziej jednorodna, a jednocześnie jak najbardziej różna od pozostałych. W zależności od zastosowanej metody, liczba grup może być determinowana w trakcie procesu uczenia modelu lub ustalana przez badacza. Liczba grup, którą ustala badacz, powinna być informowana rezultatami wcześniejszej analizy eksploracyjnej. Atrakcyjność tych metod polega na możliwości tworzenia profili grup, w ramach których badacz analizuje każdą grupę osobno i opisuje jej cechy. Dodatkowo, grupowanie otwiera nowe możliwości dalszej analizy, na przykład umożliwiając tworzenie osobnych modeli dla każdej grupy lub zastosowanie modeli hierarchicznych.

Ważnym aspektem eksploracyjnych metod analizy danych jest świadomość ich ograniczeń. Podczas EDA poszukujemy hipotez, jednak wnioski wyciągane na tym etapie są mocno ograniczone, a łatwość wyciągania błędnych wniosków stanowi jedną z głównych pułapek tego procesu. Jednym z takich błędów jest nadinterpretacja korelacji i fałszywe przypisywanie przyczynowości. Kolejnym zagrożeniem jest niewłaściwy dobór narzędzi statystyki opisowej do danej zmiennej. Innym przykładem błędu może być niewłaściwy dobór wykresu lub niepoprawne ustalenie skali na osiach wykresu, co może prowadzić do poważnych nadużyć w interpretacji danych. Wnioski o wyższej wadze wyciągane są dopiero na późniejszych etapach procesu analizy danych. EDA pełni rolę informacyjną, wskazując potencjalne hipotezy lub obszary, które wymagają dalszej analizy i weryfikacji. Na etapie EDA badacz wykonuje prace detektywistyczną na danych, poszukując wskazówek, które pozwolą nadać sens danym, zrozumieć je oraz sformułować kolejne hipotezy[[22]](#footnote-22).

**2.2 Statystyka i wnioskowanie Bayesowskie**

Współczesna analiza danych opiera się na dwóch dominujących paradygmatach statystycznych: statystyce częstościowej (en. *Frequentist statistics*) oraz statystyce bayesowskiej (en. *Bayesian statistics*). Metody statystyczne są środkami pozwalającymi na uwzględnianie błędów pomiaru oraz niepewności, dlatego główna różnica pomiędzy szkołami statystycznymi sprowadza się w istocie do odmiennej interpretacji natury i korzeni pochodzenia prawdopodobieństwa. Statystyka częstościowa, będąca przez dziesięciolecia standardowym podejściem akademickim, wywodzi się z klasycznych definicji prawdopodobieństwa formułowanych m.in. przez Jacoba Bernoulliego. W tym podejściu prawdopodobieństwo rozumiane jest jako długookresowa częstość występowania danego zdarzenia w hipotetycznie nieskończonej liczbie powtórzeń tego samego eksperymentu[[23]](#footnote-23).

Kluczową cechą statystyki częstościowej jest to, że wnioskowanie opiera się wyłącznie na tzw. *sampling probability*, czyli prawdopodobieństwie zaobserwowania konkretnych danych przy założeniu, że dana hipoteza jest prawdziwa. Odpowiedź, jaką daje ten paradygmat, brzmi: *jak bardzo nasze dane są zgodne z hipotezą zerową?*

Podejście bayesowskie proponuje odmienne podejście do interpretacji prawdopodobieństwa. Traktuje je jako miarę stopnia przekonania (subiektywnej wiarygodności), którą można aktualizować w oparciu o nowe obserwacje. W ujęciu bayesowskim interesuje nas więc nie tylko prawdopodobieństwo danych przy założeniu hipotezy, ale również prawdopodobieństwo hipotezy przy uwzględnieniu danych (en. *inferential probability*). Innymi słowy, pytanie brzmi: *jak bardzo możemy wierzyć w daną hipotezę, mając do dyspozycji określony zbiór obserwacji?*

Statystyka bayesowska zakłada, że mniej prawdopodobne (lub bardziej kontrowersyjne) hipotezy wymagają silniejszych dowodów, aby mogły zostać uznane za wiarygodne[[24]](#footnote-24). Efekt ten osiąga się poprzez zastosowanie twierdzenia Bayesa, które pozwala na łączenie wcześniejszych przekonań (rozkład a priori) z nowymi obserwacjami, prowadząc do otrzymania rozkładów a posteriori – pełnych rozkładów wartości dotyczących poszczególnych parametrów modelu. Dostęp do pełnych rozkładów jest jedną z główny zalet statystyki bayesowskiej, ponieważ umożliwia doskonałe możliwości rozumowania w warunkach niepewności.

Wielu zwolenników statystyki częstościowej wskazuje na jej obiektywność. Jednak można argumentować, że ignorowanie ustabilizowanej wiedzy naukowej (np. wyników wcześniejszych, odtwarzalnych badań) niekoniecznie czyni dane podejście bardziej obiektywnym. Statystyka bayesowska umożliwia bowiem naturalne, formalne uwzględnienie wiedzy dziedzinowej lub wyników wcześniejszych badań jako rozkładów a priori[[25]](#footnote-25).

Wnioskowanie bayesowskie to podejście statystyczne, które umożliwia aktualizację przekonań na podstawie nowych danych. Podstawowym składnikiem wnioskowania bayesowskiego jest twierdzenie Bayesa, które pozwala na odwrócenie warunkowości, czyli uzyskanie, posiadając .

Gdzie:

* : Prawdopodobieństwo zdarzenia A
* : Prawdopodobieństwo zdarzenia B
* : Prawdopodobieństwo zdarzenia A pod warunkiem B
* : Prawdopodobieństwo zdarzenia B pod warunkiem A

W kontekście wnioskowania statystycznego, zdarzenie A może reprezentować hipotezę, a zdarzenie B zaobserwowane dane[[26]](#footnote-26).

Gdzie:

* : Prawdopodobieństwo a priori — wiarygodność hipotezy przed uwzględnieniem danych
* : Wiarygodność — prawdopodobieństwo danych przy założeniu prawdziwości hipotezy
* : Dowód — całkowite prawdopodobieństwo danych, niezależnie od hipotezy
* : Prawdopodobieństwo a posteriori — zaktualizowane przekonanie o hipotezie po uwzględnieniu danych

Twierdzenie Bayesa pozwala na połączenie wierzeń dotyczących danego zjawiska z zaobserwowanymi danymi i uzyskanie miary wiarygodności tych wierzeń[[27]](#footnote-27).

Czynnik Bayesa (en. *Bayes factor*) to stosunek wiarygodności dwóch konkurujących hipotez. Porównuje on, jak dobrze każda z hipotez tłumaczy zaobserwowane dane.

Wynik ilorazu wiarygodności informuje, jak bardzo dane wspierają hipotezę ​ w porównaniu do hipotezy ​. Innymi słowy, przedstawia relatywną wiarygodność danych przy założeniu prawdziwości jednej z hipotez.

Iloraz prawdopodobieństw a priori (en. *Prior odds*), jest stosunkiem prawdopodobieństwa dwóch konkurujących hipotez przed zaobserwowaniem danych.

Odzwierciedla on początkowe przekonanie co do prawdziwości jednej hipotezy względem drugiej, zanim nastąpi obserwacja danych empirycznych. Iloraz prawdopodobieństw a priori jest kluczowy, ponieważ nie każda hipoteza jest równie prawdopodobna.

Połączenie ilorazu prawdopodobieństw a priori z czynnikiem Bayesa daje możliwość naturalnego porównywania konkurujących hipotez poprzez obliczenie ilorazu prawdopodobieństw a posteriori (en. *Posterior odds*)[[28]](#footnote-28).

Otrzymany wynik pokazuje, ile razy bardziej prawdopodobna jest hipoteza ​ niż po uwzględnieniu danych , biorąc pod uwagę zarówno wcześniejsze przekonania (iloraz prawdopodobieństw a priori), jak i zgodność danych z każdą z hipotez (czynnik Bayesa).

*Tablica 1. Wytyczne ewaluacji ilorazu prawdopodobieństw a posteriori.*

*Źródło: Opracowanie własne na podstawie Bayesian Statistics the Fun Way[[29]](#footnote-29)*

Istnieje wiele innych wytycznych dotyczących ewaluacji ilorazu prawdopodobieństw a posteriori oraz czynniku Bayesa, natomiast wszystkie służą jedynie jako zalecenia i nie są definitywnymi wyznacznikami.

Statystyka bayesowska nie ogranicza się wyłącznie do prostego zastosowania twierdzenia Bayesa czy porównywania hipotez. Jej głównym celem jest budowanie ustrukturyzowanego zrozumienia badanego zjawiska, procesu lub systemu. W tym ujęciu statystyka pełni funkcję narzędzia do konstruowania, aktualizacji i oceny modeli probabilistycznych, które odzwierciedlają rzeczywistość w sposób ilościowy i starają się wyodrębnić sygnał od szumu. W podejściu bayesowskim modele są budowane z elementarnych bloków – rozkładów prawdopodobieństwa. Rozkłady te opisują zarówno niepewność związaną z parametrami modelu (rozkład a priori), jak i mechanizm generowania danych (rozkład prawdopodobieństwa warunkowego). Struktura modelu odzwierciedla założenia na temat zależności pomiędzy obserwowanymi zmiennymi oraz procesu generującego dane[[30]](#footnote-30). Centralnym mechanizmem wnioskowania bayesowskiego jest zastosowanie twierdzenia Bayesa do aktualizacji rozkładu prawdopodobieństwa parametrów modelu w świetle nowo zaobserwowanych danych. Dzięki temu wiedza a priori zostaje skorygowana o informacje empiryczne, czego efektem jest uzyskanie rozkładu a posteriori.

Modele regresji liniowej są jednymi z najczęściej stosowanych metod wnioskowania statystycznego w nauce i praktyce biznesowej. W swojej podstawowej wersji regresja liniowa opisuje zależność pomiędzy dwiema zmiennymi ilościowymi ciągłymi: zmienną objaśniającą *x* a zmienną objaśnianą *y*. Celem jest oszacowanie kierunku i siły wpływu zmiennej *x* na wartość oczekiwaną zmiennej *y*. W podejściu bayesowskim wszystkie nieznane parametry modelu są traktowane jak zmienne losowe, dla których określane są rozkłady a priori, które reprezentują stan wiedzy o możliwych wartościach parametrów przed obserwacją danych. Następnie w oparciu o dane empiryczne, przy pomocy twierdzenia Bayesa dokonywana jest aktualizacja przekonań skutkująca uzyskaniem rozkładu a posteriori. Prosty model zakładający liniowy związek pomiędzy zmienną *x* i wartością oczekiwaną zmiennej *y* oraz normalny rozkład zmiennej *y* można przedstawić w notacji probabilistycznej, jako[[31]](#footnote-31):

Powyższy model określa generatywny proces tworzenia danych. Parametry , oraz są traktowane jako zmienne losowe, dla których należy oszacować rozkład a posteriori. Dzięki temu możliwe jest nie tylko uzyskanie punktowych estymatów parametrów, ale również pełnego rozkładu możliwych wartości. Takie podejście pozwala na bardziej przejrzyste wnioskowanie oraz uzyskanie pełnego obrazu niepewności.

Modele bayesowskie pozwalają na naturalne kodowanie założeń dotyczących parametrów oraz procesu generowania danych, co czyni je wysoce elastycznymi, w przeciwieństwie do klasycznego podejścia, w którym założenia są często sztywno narzucone. Przykładowo, w sytuacjach, gdy zmienna zależna może przyjmować wartości skrajne będące naturalną cechą badanego zjawiska, zamiast rozkładu normalnego można zastosować rozkład T-Studenta jako rozkład warunkowy. Dodatkowo, stopnie swobody rozkładu T-Studenta mogą być traktowane jako kolejny parametr modelu, dla którego również definiuje się rozkład a priori. Takie podejście pozwala modelowi automatycznie dostosować się do występowania wartości skrajnych. Kolejnym przykładem jest modelowanie heteroskedastyczności, czyli sytuacji, w której wariancja składnika losowego nie jest stała, lecz zależy od wartości jednej lub więcej zmiennych objaśniających. Odchylenie standardowe można modelować jako liniowo zależną od zmiennej lub wielu zmiennych objaśniających. Oznacza to, że nie tylko parametr średniej (np. ), ale również inne parametry rozkładu warunkowego — takie jak odchylenie standardowe — mogą być modelowane jako funkcje predyktorów[[32]](#footnote-32).

W analizie danych o strukturze grupowej (np. dane z podziałem na klasy, regiony, jednostki organizacyjne) istnieją trzy główne podejścia do modelowania statystycznego:

1. Całkowite łączenie grup (en. *Complete pooling*)

W tym podejściu obserwacje ze wszystkich grup są traktowane jako jednorodna całość, a informacja o przynależności grupowej jest ignorowana. Tworzony jest jeden globalny model, wspólny dla wszystkich obserwacji. W niektórych przypadkach całkowite łączenie grup może prowadzić do utraty istotnych informacji, szczególnie gdy występują różnice między grupami.

1. Brak łączenia grup (en. *No pooling*)

Każda grupa traktowana jest jako niezależna. Dla każdej z nich estymowane są osobne parametry modelu, zazwyczaj z wykorzystaniem niezależnych rozkładów a priori. Podejście to pozwala na uchwycenie różnic między grupami, jednak przy niewielkiej liczbie obserwacji w danej grupie prowadzi do niestabilnych estymatorów.

1. Częściowe łączenie grup (en. *Partial pooling*)

Stanowi kompromis pomiędzy powyższymi podejściami. Zakłada się, że grupy są powiązane i dzielą pewne wspólne właściwości, które modelowane są poprzez tzw. *hyperprior*, czyli globalny rozkład a priori wyższego poziomu. Każda grupa ma swoje parametry, ale są one warunkowane przez wspólny, globalny rozkład a priori.

Trzecie podejście, nazywane również modelami hierarchicznymi, pozwala na uzyskanie efektu skurczenia (en. *Shrinkage*), polegający na „przyciąganiu” estymatów parametrów grupowych w kierunku wartości globalnych. Efekt skurczania jest szczególnie przydatny w sytuacjach, gdy jedna z grup posiada znacząco mniej obserwacji. Skurczanie pomaga również przeciwdziałać efektowi nadmiernego dopasowania (en. *Overfitting*). Modele hierarchiczne pozwalają na jednoczesne uchwycenie efektu globalnego (wspólnego trendu dla wszystkich grup) oraz efektów specyficznych dla poszczególnych grup[[33]](#footnote-33).

**3 Analiza Danych**

Analiza własna została przeprowadzona w dwóch powiązanych etapach, odzwierciedlających logiczną strukturę podejścia do analizy danych. W pierwszej części zastosowano eksploracyjną analizę danych (EDA), której celem było uzyskanie wstępnego wglądu w strukturę i charakterystykę zmiennych zawartych w zbiorze danych. Analiza ta pozwoliła na identyfikację potencjalnych zależności, rozkładów zmiennych, oraz problemów, takich jak brakujące dane, obserwacje skrajne czy zmienność między jednostkami.

W kolejnej części zaprezentowano formalne modelowanie statystyczne z wykorzystaniem podejścia bayesowskiego. Przyjęto różne założenia dotyczące struktury danych, w tym modeli hierarchicznych, co umożliwiło uwzględnienie efektów grupowych oraz niepewności związanej z estymacją parametrów. Modele te zostały ocenione zarówno pod względem dopasowania, jak i interpretowalności wyników. Ostatnia część rozdziału koncentruje się na ocenie wiarygodności oraz na interpretacji wyników w kontekście analizowanego zjawiska.

Wykorzystywany zbiór danych obejmuje dane z urządzeń ubieralnych noszonych przez piętnaście pielęgniarek, które monitorowały m.in. tętno, reakcję elektrodermalną (*en. Electrodermal activity, EDA*) oraz temperaturę skóry.

**3.1 Eksploracyjna analiza danych (EDA)**

Podczas eksploracji danych duży nacisk położono na procesowanie danych przygotowujących je do dalszej analizy. Dokonano ponownego próbkowania danych do interwałów czasowych wynoszących 1 sekundę, co pozwoliło na znaczną redukcję rozmiaru zbioru danych z pierwotnych 11,5 miliona wierszy do 356 561 wierszy, przy jednoczesnym zachowaniu informacji umożliwiających eksploracyjną analizę danych.

W kolejnym kroku zbiór danych został sprawdzony pod kątem brakujących wartości. Stwierdzono, że w zbiorze danych nie występują brakujące wartości.

Stworzono tablicę statystyk opisowych.

*Tablica 2. Statystyki opisowe dla zmiennych ilościowych.*

*Źródło: Opracowanie własne*

Następnie skupiono się na wizualizacji danych. Stworzono wykresy pudełkowe, macierz korelacji oraz wykresy rozrzutu dla cech ciągłych zbioru danych.

*A diagram of a graph

AI-generated content may be incorrect.Wykres 8. Rozkłady zmiennych ilościowych w podziale na uczestników badania.*

*Źródło: Opracowanie własne*

Wykres obrazuje znaczną zmienność wartości pomiarowych pomiędzy poszczególnymi uczestnikami badania. W szczególności obiekt E4 charakteryzuje się dużymi dla zmiennej opisującej reakcję elektrodermalną (EDA). Ponadto zaobserwowano znaczną różnorodność wartości EDA na poziomie indywidualnych uczestników.

Podobne zróżnicowanie dotyczyło zmiennej opisującej tętno. Średnia arytmetyczna tętna dla uczestnika z najniższymi odczytami wynosiła 75,47 uderzeń na minutę, natomiast dla osoby z najwyższą średnią – 84,68 uderzeń na minutę.

Temperatura skóry również różnicowała badanych. Można zaobserwować dwie główne grupy obiektów, w kontekście miar położenia temperatury. Pierwsza z nich zawiera niższe wartości, w okolicach 29 stopni Celsjusza, a druga wyższe, około 33 stopni. Szczególną uwagę zwracają wartości też niskie wartości zmiennej oscylujące wokół 25°C. Należy jednak pamiętać, że odczyt sensorów różni się od tradycyjnych metod pomiaru temperatury ciała. Co więcej, urządzenia ubieralne mierzyły temperaturę skóry na nadgarstku, który jest bardziej podatny na wychłodzenie niż wnętrze organizmu.

Uwagę zwraca również duża liczba wartości skrajnych, zidentyfikowanych na podstawie rozstępu międzykwartylowego, we wszystkich badanych cechach. Na tym etapie analizy nie można jednak jednoznacznie stwierdzić, że są to błędne odczyty - wszystkie wartości mieszczą się w przewidywalnych zakresach zmienności.

Następnie stworzono macierz korelacji, która dostarcza informacji o współzależności zmiennych.

*A yellow and purple squares with white text

AI-generated content may be incorrect.Wykres 9. Macierz korelacji dla zmiennych ilościowych*

*Źródło: Opracowanie własne*

Analiza korelacji wskazuje, że najsilniejszy związek występuje pomiędzy temperaturą skóry a aktywnością elektrodermalną (r = 0,35). Oznacza to umiarkowaną dodatnią korelację – wyższe wartości EDA są związane z wyższą temperaturą skóry. Z kolei korelacja pomiędzy tętnem a temperaturą skóry jest słabsza (r = 0,16), ale nadal dodatnia, co sugeruje, że wyższa temperatura skóry może być nieznacznie związana z wyższym tętnem. Najmniejsza współzależność dodatnia występuje pomiędzy EDA a tętnem (r = 0,14).

Siła związku zmienia się po stratyfikacji ze względu na deklarowany poziom stresu.

A chart of different colors

AI-generated content may be incorrect.*Wykres 10. Macierz korelacji dla zmiennych ilościowych z podziałem na poziom stresu.*

*Źródło: Opracowanie własne*

Przykładowo, wśród badanych deklarujących niski/średni poziom stresu korelacja pomiędzy temperaturą skóry a EDA jest silniejsza (r = 0.5).

W kolejnym kroku skonstruowano wykresy rozrzutu dla każdej pary zmiennych ilościowych. Dodatkowo obserwowany był rozkład każdej cechy przy użyciu wykresu jądrowego estymatora gęstości, czyli gładkiego histogramu. Cała wizualizacja jest stratyfikowana ze względu na poziom stresu.

A group of graphs showing different colored dots

AI-generated content may be incorrect.*Wykres 11. Wykresy rozrzutu i rozkłady zmiennych ilościowych z podziałem na poziom stresu.*

*Źródło: Opracowanie własne*

Pierwszym wnioskiem z wykresu, jest zauważalna nierównowaga w liczbie obserwacji w poszczególnych kategoriach stresu. Partycypanci najwięcej czasu spędzali odczuwając wysoki poziomu stresu.

Widoczna jest duża asymetria prawostronna w rozkładzie zmiennej opisującej aktywność elektrodermalną. Rozkład tętna jest najbardziej zbliżony do rozkładu normalnego. W przypadku temperatury skóry, ponownie można zaobserwować dwie grupy obiektów.

Eksploracyjna analiza danych dostarczyła cennych informacji o związkach pomiędzy zmiennymi, kształtach rozkładów zmiennych, różnic fizjologicznych pomiędzy poszczególnymi obiektami oraz dostarczyła wskazówki dotyczące kierunku dalszej analizy.

**3.2 Wpływ deklarowanego poziomu stresu na zmienne fizjologiczne**

W pierwszej kolejności uwagę skupiono na zmianach i różnicach w zmierzonych zmiennych fizjogicznych wśród partycypantów. Konkretnie, usiłowano odpowiedzieć na pytanie badawcze: Czy istnieje, i jaka jest siła efektu deklarowanego stresu na zmienne fizjologiczne. Zmienne fizjologiczne istotnie różnią się w zależności od obiektu, więc skupiono się na hierarchicznych model liniowych, dla trzech zmiennych objaśnianych: Temperatura skóry, tętno oraz aktywność elektrodermalna.

W celu wybrania odpowiedniego modelu dla każdej zmiennej dopasowano cztery specyfikacje modeli. Zastosowano ten krok, aby pomóc w wyborze najodpowiedniejszych modeli dla każdej zmiennej zależnej. Testowano cztery różne podejścia:

1. Dwa modele regresji liniowej bez łączenia pomiędzy grupami. W każdym przypadku każda pielęgniarka otrzymuje oddzielne, niezależne rozkłady a priori dla parametrów α, β oraz σ.
   * Pierwszy model zakłada monotoniczną, liniową relację na każdym poziomie stresu (0, 1, 2).
   * Drugi model korzysta z kodowania zerojedynkowego dla zmiennej stresu, co pozwala na bardziej elastyczną estymację parametrów β.
2. Hierarchiczny model liniowy, który pozwala na uwzględnienie efektu skurczania parametrów.
3. Model liniowy zakładający zmienność wariancji, twierdzący, że wariancja (σ) zmiennej zależnej zależy od poziomu stresu. Logarytm parametru σ był modelowany jako funkcja liniowa w odniesieniu do poziomu stresu dla każdej pielęgniarki.

Gdzie:

* i: indeks pielęgniarki,
* j: indeks obserwacji,
* αᵢ: wyraz wolny dla pielęgniarki,
* βᵢ: współczynnik regresji (efekt stresu) dla pielęgniarki i,
* σᵢ: odchylenie standardowe dla pielęgniarki i,
* ∈ {0, 1, 2}: poziom stresu (0 = Brak, 1 = Niski/Średni, 2 = Wysoki).
* : zmienna zależna (Temperatura skóry, Tętno, Aktywność elektrodermalna)

Gdzie:

* i: indeks pielęgniarki,
* j: indeks obserwacji,
* αᵢ: wyraz wolny dla pielęgniarki i,
* β₀: współczynnik regresji dla poziomu stresu "Niski/Średni" dla pielęgniarki i,
* β₁: współczynnik regresji dla poziomu stresu "Wysoki" dla pielęgniarki i,
* σᵢ: odchylenie standardowe dla pielęgniarki i,
* NiskiŚrᵢⱼ ∈ {0, 1}: zmienna binarna wskazująca, czy poziom stresu w obserwacji j pielęgniarki i to "Niski/Średni" (1 = Tak, 0 = Nie),
* Wysokiᵢⱼ ∈ {0, 1}: zmienna binarna wskazująca, czy poziom stresu w obserwacji j pielęgniarki i to "Wysoki" (1 = Tak, 0 = Nie),
* yᵢⱼ: zmienna zależna (Temperatura skóry, Tętno, Aktywność elektrodermalna).

Gdzie:

* i: indeks pielęgniarki
* j: indeks obserwacji
* αᵢ: wyraz wolny dla pielęgniarki i
* βᵢ: współczynnik efektu stresu dla pielęgniarki i
* σᵢ: odchylenie standardowe dla pielęgniarki i
* stresᵢⱼ: stres w obserwacji j dla pielęgniarki i
* yᵢⱼ: zmienna zależna (Temperatura skóry, Tętno, Aktywność elektrodermalna)

Gdzie:

* i: indeks pielęgniarki
* j: indeks obserwacji
* αᵢ: wyraz wolny dla pielęgniarki i
* β₀ᵢ: współczynnik regresji dla poziomu stresu "Niski/Średni" dla pielęgniarki i
* β₁ᵢ: współczynnik regresji dla poziomu stresu "Wysoki" dla pielęgniarki i
* σₐᵢ: punkt przecięcia dla log(sigma) dla pielęgniarki i
* σᵦᵢ:: nachylenie dla log(sigma) dla pielęgniarki i
* ∈ {0, 1, 2}: poziom stresu (0 = Brak, 1 = Niski/Średni, 2 = Wysoki).
* σᵢ: odchylenie standardowe dla pielęgniarki i
* yᵢⱼ: zmienna zależna (Temperatura skóry, Tętno, Aktywność elektrodermalna)

W celu wyboru ostatecznego zestawu modeli stworzono tablicę porównującą wymienione podejścia, używając PSIS-LOO-CV (en. *Pareto-smoothed importance sampling leave-one-out cross-validation*). Dla każdego modelu obliczono wartość ELPD-LOO (en. *Expected Log Predictive Density*) oraz wygenerowano wykres trajektorii łańcucha MCMC (en. *Trace plot*), w celuweryfikacji niezawodności uzyskanych rozkładów a posteriori.

Ostatecznie, zdecydowano się na zastosowanie modelu, który w pewnym sensie łączy rozwiązania używane w modelach w porównaniach. Model zakładający heteroskedastyczność posiadał najwyższy wskaźnik ELPD-LOO dla każdej zmiennej, sugerując zasadność tego podejścia. Rozkłady a posteriori parametrów modeli bez struktury hierarchicznej, sugerują, że wprowadzenie skurczenia powinno wspomóc interpretowalność modeli. W związku z tym ostatecznie wybrano hierarchiczny model liniowy pozwalający na łączenie informacji pomiędzy grupami, zakładający zmienność wariancji w zależności od poziomu stresu. Dodatkowo wprowadzono alternatywne funkcje wiarygodności dla zmiennych zależnych, które lepiej pasują do ich empirycznych rozkładów.

Gdzie:

* i: indeks pielęgniarki
* j: indeks obserwacji
* αᵢ: wyraz wolny dla pielęgniarki i
* β₀ᵢ: współczynnik regresji dla poziomu stresu "Niski/Średni" dla pielęgniarki i
* β₁ᵢ: współczynnik regresji dla poziomu stresu "Wysoki" dla pielęgniarki i
* σₐᵢ: punkt przecięcia dla log(sigma) dla pielęgniarki i
* σᵦᵢ:: nachylenie dla log(sigma) dla pielęgniarki i
* ∈ {0, 1, 2}: poziom stresu (0 = Brak, 1 = Niski/Średni, 2 = Wysoki).
* σᵢ: odchylenie standardowe dla pielęgniarki i
* yᵢⱼ: zmienna zależna (Temperatura skóry, Tętno, Aktywność elektrodermalna)

1. Gary Wolf, „Know Thyself: Tracking Every Facet of Life, from Sleep to Mood to Pain, 24/7/365”, *Wired*, dostęp 12 listopad 2024, https://www.wired.com/2009/06/lbnp-knowthyself/. [↑](#footnote-ref-1)
2. Deborah Lupton, *The Quantified Self: A Sociology of Self-Tracking* (Cambridge, UK: Polity, 2016). [↑](#footnote-ref-2)
3. Steve Mann, „Wearable computing as means for personal empowerment”, w *Proc. 3rd Int. Conf. on Wearable Computing (ICWC)*, 1998, 51–59, https://www.researchgate.net/profile/Samuel-Mann-2/publication/235220249\_Wearable\_computing\_as\_a\_means\_for\_personal\_empowerment/links/00b495321165583304000000/Wearable-computing-as-a-means-for-personal-empowerment.pdf. [↑](#footnote-ref-3)
4. Steve Mann, „Mediated Reality”, *Linux Journal* 1999, nr 59es (1 marzec 1999): 5-es. [↑](#footnote-ref-4)
5. „Hacking Health: How Consumers Use Smartphones and Wearable Tech to Track Their Health”, Nielsen, dostęp 27 listopad 2024, https://www.nielsen.com/insights/2014/hacking-health-how-consumers-use-smartphones-and-wearable-tech-to-track-their-health/. [↑](#footnote-ref-5)
6. Eun Kyoung Choe i in., „Understanding Quantified-Selfers’ Practices in Collecting and Exploring Personal Data”, w *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI ’14: CHI Conference on Human Factors in Computing Systems, Toronto Ontario Canada: ACM, 2014), 1143–52, https://doi.org/10.1145/2556288.2557372. [↑](#footnote-ref-6)
7. André Henriksen i in., „Dataset of Fitness Trackers and Smartwatches to Measuring Physical Activity in Research”, *BMC Research Notes* 15, nr 1 (16 lipiec 2022): 258, https://doi.org/10.1186/s13104-022-06146-5. [↑](#footnote-ref-7)
8. „What Is Player Load?”, Catapult Support, 31 styczeń 2024, https://support.catapultsports.com/hc/en-us/articles/360000510795-What-is-Player-Load. [↑](#footnote-ref-8)
9. *Wearable technologies and sports analytics, University of Michigan* (Coursera, b.d.), dostęp 31 grudzień 2024. [↑](#footnote-ref-9)
10. Tim J Gabbett, „The training—injury prevention paradox: should athletes be training smarter and harder?”, *British Journal of Sports Medicine* 50, nr 5 (marzec 2016): 273–80, https://doi.org/10.1136/bjsports-2015-095788. [↑](#footnote-ref-10)
11. Leandro Carbone i in., „Is the Relationship between Acute and Chronic Workload a Valid Predictive Injury Tool? A Bayesian Analysis”, *Journal of Clinical Medicine* 11, nr 19 (styczeń 2022): 5945, https://doi.org/10.3390/jcm11195945. [↑](#footnote-ref-11)
12. Matthew T. O. Worsey i in., „An Evaluation of Wearable Inertial Sensor Configuration and Supervised Machine Learning Models for Automatic Punch Classification in Boxing”, *IoT* 1, nr 2 (grudzień 2020): 360–81, https://doi.org/10.3390/iot1020021. [↑](#footnote-ref-12)
13. Dave Ebbelaar, „Exploring the Possibilities of Context-Aware Applications for Strength Training” (Amsterdam, Vrije Universiteit), dostęp 17 grudzień 2024, https://github.com/daveebbelaar/tracking-barbell-exercises/tree/master. [↑](#footnote-ref-13)
14. Melisa Junata i Raymond Kai-Yu Tong, „Chapter 1 - Wearable Technology in Medicine and Health Care: Introduction”, w *Wearable Technology in Medicine and Health Care*, red. Raymond Kai-Yu Tong (Academic Press, 2018), 1–5, https://doi.org/10.1016/B978-0-12-811810-8.00001-4. [↑](#footnote-ref-14)
15. (PubMed), dostęp 31 grudzień 2024, https://pubmed.ncbi.nlm.nih.gov/?term=%28wearables%29+AND+%28%28health+care%29+OR+%28medicine%29%29&filter=years.2005-2025. [↑](#footnote-ref-15)
16. Junata i Tong, „Chapter 1 - Wearable Technology in Medicine and Health Care”. [↑](#footnote-ref-16)
17. Patrick Slevin i Brian Caulfield, „Chapter 13 - Patient-Generated Health Data: Looking Toward Future Health Care”, w *Wearable Technology in Medicine and Health Care*, red. Raymond Kai-Yu Tong (Academic Press, 2018), 261–73, https://doi.org/10.1016/B978-0-12-811810-8.00013-0. [↑](#footnote-ref-17)
18. Glenn J. Myatt i Wayne P. Johnson, *Making Sense of Data I: A Practical Guide to Exploratory Data Analysis and Data Mining*, Second edition, Safari Tech Books Online (Hoboken, New Jersey: Wiley, 2014). [↑](#footnote-ref-18)
19. Mark Hoogendoorn i Burkhardt Funk, *Machine Learning for the Quantified Self: On the Art of Learning from Sensory Data*, Cognitive Systems Monographs 35 (Cham: Springer, 2018), https://doi.org/10.1007/978-3-319-66308-1. [↑](#footnote-ref-19)
20. N. N. R. Ranga Suri, *Outlier Detection: A Data Mining Perspective*, Intelligent Systems Reference Library, v. 155 (Cham: Springer, 2019). [↑](#footnote-ref-20)
21. Brett Kennedy, *Outlier Detection in Python*, 1st ed (New York: Manning Publications Co. LLC, 2025). [↑](#footnote-ref-21)
22. John Wilder Tukey, *Exploratory Data Analysis*, Addison-Wesley Series in Behavioral Science (Reading (Mass.) Menlo Park (Calif.) London [etc.]: Addison-Wesley publ, 1977). [↑](#footnote-ref-22)
23. Aubrey Clayton, *Bernoulli’s Fallacy: Statistical Illogic and the Crisis of Modern Science* (New York: Columbia University Press, 2021). [↑](#footnote-ref-23)
24. Clayton. [↑](#footnote-ref-24)
25. Osvaldo Martin, *Bayesian Analysis with Python: A Practical Guide to Probabilistic Modeling*, 3. wyd. (Birmingham: Packt Publishing Limited, 2024). [↑](#footnote-ref-25)
26. Will Kurt, *Bayesian Statistics the Fun Way: Understanding Statistics and Probability with Star Wars, LEGO, and Rubber Ducks* (San Francisco, CA: No Starch Press, 2019). [↑](#footnote-ref-26)
27. Kurt. [↑](#footnote-ref-27)
28. Kurt. [↑](#footnote-ref-28)
29. Kurt. [↑](#footnote-ref-29)
30. Martin, *Bayesian Analysis with Python*. [↑](#footnote-ref-30)
31. Martin. [↑](#footnote-ref-31)
32. Martin. [↑](#footnote-ref-32)
33. Richard McElreath, *Statistical rethinking: a Bayesian course with examples in R and Stan*, Second edition, Chapman & Hall/CRC texts in statistical science series (Boca Raton: CRC Press, 2020). [↑](#footnote-ref-33)