



UNIwersYTET GDAŃSKI

Jakub Gorgoń
Nr albumu: 286131

Quantified Self, technologie ubieralne i wnioskowanie bayesowskie w modelowaniu fizjologicznych reakcji na stres

Kierunek studiów: Informatyka i ekonometria
Poziom studiów: I stopnia - licencjackie
Profil studiów: ogólnoakademicki
Forma studiów: studia stacjonarne

Praca dyplomowa licencjacka
wykonana pod kierunkiem
dr Anna Gierusz-Matkowska

Gdańsk, 2025 r.

Spis treści

Wstęp.....	2
Rozdział 1. Idea <i>quantified self</i> : Zbieranie i analiza danych osobistych	4
1.1 Wprowadzenie i kluczowe zagadnienia	4
1.2 Powody i motywacje do zbierania i analizy danych osobistych.....	6
1.3 Technologie i urządzenia ubieralne	7
1.3.1 Sensory w urządzeniach ubieralnych	7
1.3.2 Aplikacje urządzeń ubieralnych.....	9
Rozdział 2. Metodologia	15
2.1 Eksploracyjna analiza danych (EDA).....	15
2.2 Statystyka i wnioskowanie bayesowskie	19
Rozdział 3. Modelowanie wpływu stresu na temperaturę i tętno na podstawie danych z urządzeń ubieralnych pielęgniarek	28
3.1 Eksploracyjna analiza badanego zbioru danych	28
3.2 Wpływ deklarowanego poziomu stresu na zmienne fizjologiczne	32
Zakończenie	45
Bibliografia.....	46

Wstęp

W ostatnich latach obserwuje się dynamiczny rozwój idei *Quantified Self*, której istotą jest systematyczne gromadzenie oraz analiza danych dotyczących własnego ciała i codziennego funkcjonowania. Trend ten obejmuje wykorzystanie urządzeń i aplikacji służących do monitorowania aktywności fizycznej, parametrów fizjologicznych, snu, nastroju czy poziomu stresu. Popularyzacja technologii ubieralnych znacząco zwiększyła możliwości zbierania danych osobistych w sposób ciągły, pasywny i nieinwazyjny. Zjawisko to znajduje zastosowanie zarówno w sporcie, opiece zdrowotnej, jak i w codziennym dążeniu do poprawy samopoczucia, zdrowia oraz efektywności funkcjonowania.

Celem niniejszej pracy jest określenie wpływu poziomu stresu na reakcje fizjologiczne organizmu przy wykorzystaniu modelowania bayesowskiego oraz danych zgromadzonych przez urządzenia ubieralne. Pytania badawcze postawione w pracy to:

- czy dane gromadzone przez urządzenia ubieralne mogą wspierać analizę fizjologicznych reakcji na stres?
- oraz w jaki sposób reakcje fizjologiczne różnią się między poziomami stresu?

W analizie wykorzystano dane wtórne, zebrane od personelu pielęgniarstwa w szpitalu, obejmujące zmienne fizjologiczne oraz deklarowany poziom stresu. Analiza danych przebiegała dwuetapowo. W pierwszym etapie przeprowadzono eksploracyjną analizę danych (EDA), której celem była wizualizacja zmiennych, badanie ich rozkładów, analiza współzależności oraz podstawowych statystyk opisowych. W drugim etapie wykorzystano podejście bayesowskie, do oszacowania wpływu deklarowanego stresu na zmienne fizjologiczne. Zastosowanie hierarchicznych liniowych modeli bayesowskich umożliwiło nie tylko uwzględnienie dodatkowych czynników wpływających na niepewność estymacji, ale także modelowanie różnic pomiędzy poszczególnymi uczestnikami badania.

W rozdziale pierwszym opisano genezę idei *Quantified Self* oraz opisano główne czynniki, które doprowadziły do popularyzacji tego zjawiska. Szczególną uwagę poświęcono sensorom wykorzystywanym w urządzeniach ubieralnych oraz ich zastosowaniom w sporcie i opiece zdrowotnej.

Rozdział drugi zawiera opis wykorzystanych metod ilościowych — od eksploracyjnej analizy danych po wnioskowanie bayesowskie. Uwzględniono zarówno techniczne, jak i epistemologiczne aspekty w podejściach do statystyki i analizy danych.

W rozdziale trzecim zaprezentowano dane dotyczące zmiennych fizjologicznych i deklarowanego poziomu stresu, a następnie omówiono wyniki modeli statystycznych badających wpływ stresu na reakcje fizjologiczne w badanej grupie uczestników.

Rozdział 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 nastrój¹. 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 popularności².

Wielu badaczy, inżynierów i informatyków eksperymentowało z technologiami cyfrowymi, w szczególności komputerami ubieralnymi (ang. *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 skupienia³.

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

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/>.

² Deborah Lupton, *The Quantified Self: A Sociology of Self-Tracking*, (Cambridge: Polity, 2016).

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.

monitorowanie; wyobrażał sobie technologię ubieralną jako sposób na stworzenie tego, co nazwał „rzeczywistością modyfikowaną” (ang. „*mediated reality*”) – zdolność do zmiany lub ulepszania doświadczeń sensorycznych poprzez cyfrową augmentację⁴. 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 wartości 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

⁴ Steve Mann, „Mediated Reality” *Linux Journal*, 1999, nr 59es, (1 marzec 1999), 5-es.

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 rocznie⁵.

W innym badaniu naukowcy z Uniwersytetu Waszyngtońskiego i Microsoftu przyjrzeni 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 nastrój⁶. Motywacje do samodzielnego zbierania i analizy danych można podzielić na trzy główne kategorie:

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/>.

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: ACM, 2014): 1143–52, <https://doi.org/10.1145/2556288.2557372>.

- Poprawa zdrowia: zarządzanie lub leczenie stanów chorobowych, ale również osiągnięcie 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

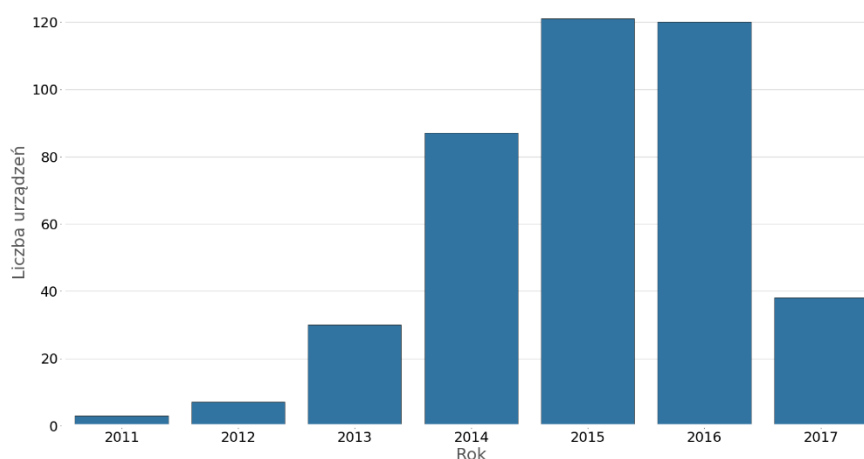
Sensory wykorzystywane w urządzeniach ubieralnych można podzielić na wewnętrzne i zewnętrzne w zależności od rodzaju zbieranych danych. Sensory zewnętrzne są umieszczone na powierzchni ciała, ale nie wymagają bezpośredniego kontaktu ze skórą i mierzą zjawiska środowiskowe, które zachodzą poza ciałem 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 praktycznych⁷.

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”.

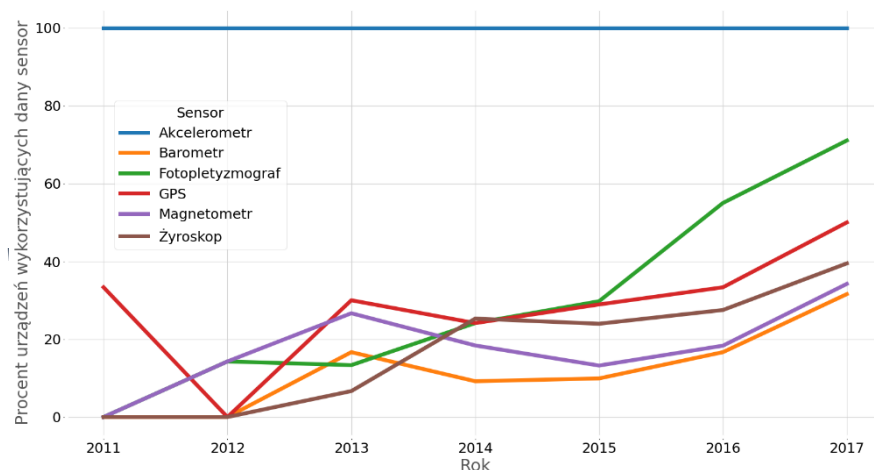
Można zaobserwować wyraźny wzrost liczby wprowadzanych na rynek urządzeń, z wyjątkiem okresu 2016-2017, co może być wynikiem ograniczenia zbierania danych do połowy 2017 roku.

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>.

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 sportowców-amatorów 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” (ang. „*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 zawodnika⁸

$$\text{Skumulowane obciążenie zawodnika} = \sum_{t=0}^{t=n} \sqrt{(Z_{t+1} - Z_t)^2 + (X_{t+1} - X_t)^2 + (Y_{t+1} - Y_t)^2}$$

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

⁸ „What Is Player Load?” Catapult Support 31 styczeń 2024, <https://support.catapultsports.com/hc/en-us/articles/360000510795-What-is-Player-Load>.

obciążenia zawodnika w krótkim okresie do obciążenia w długim okresie. Takie obliczenie jest znane jako proporcja obciążenia gwałtownego do przewlekłego (ang. *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 tygodni⁹.

$$ACWR = \frac{\text{Obciążenie gwałtowne}(t)}{\text{Obciążenie przewlekłe}(t)}$$

gdzie:

- $\text{Obciążenie gwałtowne}(t) = \sum_{i=t-T_{\text{gwałtowne}}}^t \text{Obciążenie}(i)$,
- $\text{Obciążenie przewlekłe}(t) = \sum_{i=t-T_{\text{przewlekłe}}}^t \text{Obciążenie}(i)$,
- $T_{\text{gwałtowne}}$ to długość okna czasowego dla obciążenia gwałtownego,
- $T_{\text{przewlekłe}}$ to długość okna czasowego dla obciążenia przewlekłego.

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 okresie¹⁰.

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 zewnętrzne¹¹.

⁹ *Wearable technologies and sports analytics*, University of Michigan t. Sports Performance Analytics Coursera, (Michigan: Coursera, 2021).

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>.

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>.

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 sporcie¹².

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 personalnych¹³.

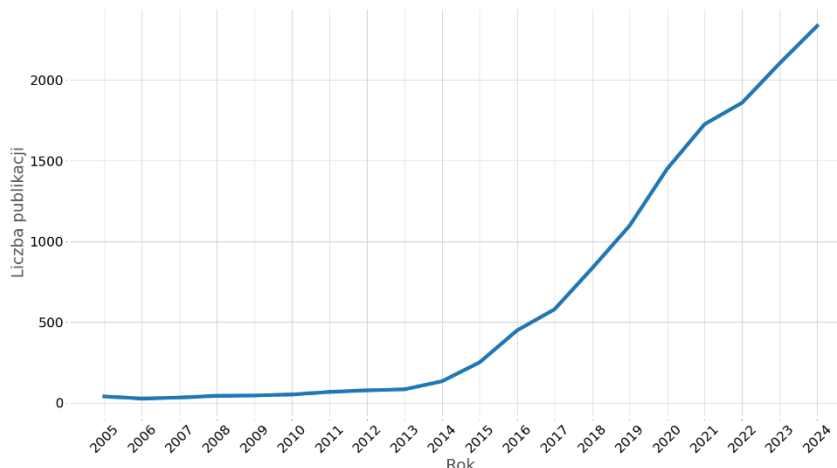
Kolejnym zastosowaniem urządzeń ubieralnych jest medycyna i służba zdrowia. Urządzenia te pozwalają na ciągłe monitorowanie oznak fizjologicznych pacjentów i stają się coraz dokładniejsze w pomiarach. Globalne trendy oraz liczba publikacji naukowych wskazują na

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>.

¹³ Dave Ebbelaar, „Exploring the Possibilities of Context-Aware Applications for Strength Training”, (Amsterdam, Vrije Universiteit, 2019), <https://github.com/daveebelaar/tracking-barbell-exercises/tree/master>.

gwałtowny rozwój wykorzystania urządzeń ubieralnych w medycynie, sugerując, że w przyszłości staną się one istotnym elementem systemu opieki zdrowotnej¹⁴.

Wykres 3. Liczba publikacji na temat urządzeń ubieralnych w kontekście opieki zdrowotnej i medycyny na PubMed (2005–2024)



Źródło: Opracowanie własne na podstawie danych z PubMed¹⁵

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 lekarzami¹⁶.

PGHD (ang. *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

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>.

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>.

¹⁶ 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>.

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 ciągle 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 leczenia¹⁷.

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>.

Rozdział 2. Metodologia

2.1 Eksploracyjna analiza danych (EDA)

Eksploracyjna analiza danych (*ang. 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 grupowanie 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 kluczowa, ponieważ wykorzystuje ludzką zdolność do identyfikacji skomplikowanych wzorców w formach graficznych¹⁸. 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 (trudniej 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

¹⁸ Glenn J. Myatt i Wayne P. Johnson, *Making Sense of Data I: A Practical Guide to Exploratory Data Analysis and Data Mining* 2. wyd. Safari Tech Books Online, (Hoboken: Wiley, 2014).

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. Próba polepszenia imputacji miarą tendencji centralnej może być zawężanie kryterium obliczania wartości. Przykładem jest wykorzystanie innej zmiennej jakościowej w celu podziału zbioru danych na grupy i na tej podstawie obliczenie 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 dla jednej zmiennej. W takim przypadku można traktować ją jako zmienną zależną (Y) i opracować model, który estymuje brakujące wartości na podstawie pozostałych dostępnych danych¹⁹.

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 wyniki analizy, prowadzić do błędnych wniosków i ostatecznie nietrafionych decyzji²⁰. 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 kwartyłu (Q1) lub powyżej trzeciego kwartyłu (Q3). Matematycznie granice wykrywania wartości skrajnych można wyrazić jako²¹:

$$\text{Dolna granica} = Q_1 - 1.5 \times IQR$$

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>.

²⁰ N. N. R. Ranga Suri, *Outlier Detection: A Data Mining Perspective* Intelligent Systems Reference Library, v. 155, (Cham: Springer, 2019).

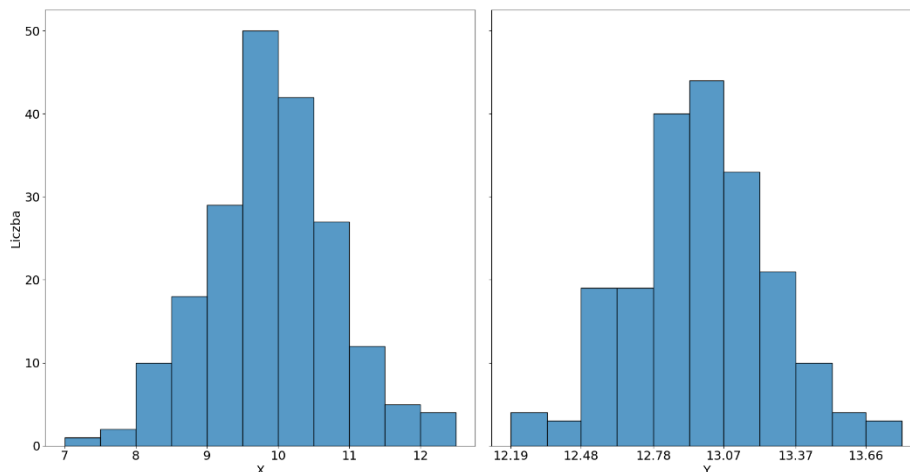
²¹ Brett Kennedy, *Outlier Detection in Python* 1. wyd., (New York: Manning Publications Co. LLC, 2025).

$$\text{Górna granica} = Q_3 + 1.5 \times IQR$$

$$\text{gdzie } IQR = Q_3 - Q_1$$

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 nietypowych. Wykresy rozrzutu pomagają w identyfikacji wartości skrajnych w kontekście dwóch zmiennych. Wykresy 4 i 5 obrazują potencjał 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.

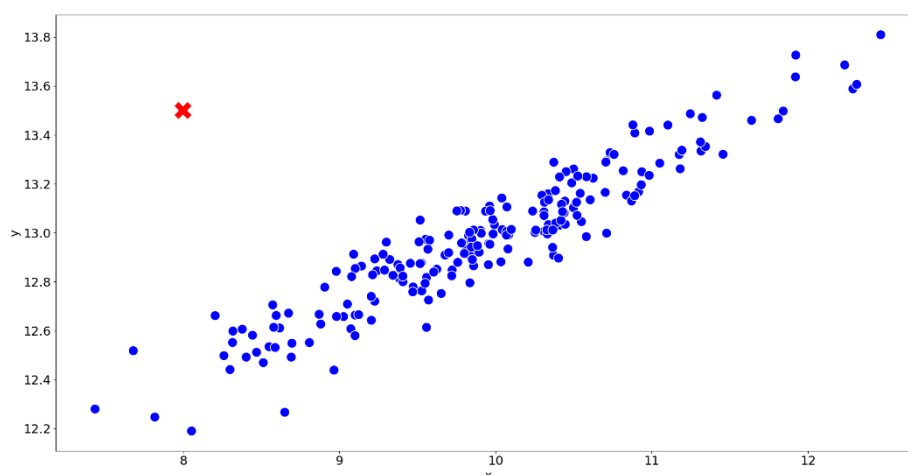
Wykres 4. Rozkład syntetycznych zmiennych x i y



Źródło: Opracowanie własne.

Zmienne analizowane osobno nie wykazują znaków występowania wartości skrajnych. Dopiero po wizualizacji 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 5. 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 zidentyfikowanie 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. Zaletą usunięcia takiej wartości jest potencjalne lepsze przygotowanie danych do 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 wynikać z rezultatów 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 hipotezy²².

2.2 Statystyka i wnioskowanie bayesowskie

Współczesna analiza danych opiera się na dwóch dominujących paradygmatach statystycznych: statystyce częstościowej (ang. *Frequentist statistics*) oraz statystyce bayesowskiej (ang. *Bayesian statistics*). Metody statystyczne są środkami pozwalającymi na uwzględnianie błędów pomiaru oraz niepewności, dlatego główna różnica pomiędzy szkołami

²² John Wilder Tukey, *Exploratory Data Analysis* Addison-Wesley Series in Behavioral Science, (Reading: Addison-Wesley publ, 1977).

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 eksperymentu²³.

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 (ang. *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 wiarygodne²⁴. 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 głównych zalet statystyki bayesowskiej, ponieważ umożliwia doskonale 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

²³ Aubrey Clayton, *Bernoulli's Fallacy: Statistical Illogic and the Crisis of Modern Science*, (New York: Columbia University Press, 2021).

²⁴ Aubrey Clayton, *Bernoulli's Fallacy: Statistical Illogic and the Crisis of Modern Science*, (New York: Columbia University Press, 2021).

obiektywnym. Statystyka bayesowska umożliwia bowiem naturalne, formalne uwzględnienie wiedzy dziedzinowej lub wyników wcześniejszych badań jako rozkładów a priori²⁵.

Wnioskowanie bayesowskie to podejście statystyczne, które umożliwia aktualizację przekonań na podstawie nowych danych. Podstawą wnioskowania bayesowskiego jest twierdzenie Bayesa, które pozwala na odwrócenie warunkowości, czyli uzyskanie $P(A | B)$, posiadając $P(B | A)$.

$$P(A | B) = \frac{P(A) \cdot P(B | A)}{P(B)}$$

gdzie:

- $P(A)$: prawdopodobieństwo zdarzenia A,
- $P(B)$: prawdopodobieństwo zdarzenia B,
- $P(A | B)$: prawdopodobieństwo zdarzenia A pod warunkiem B,
- $P(B | A)$: prawdopodobieństwo zdarzenia B pod warunkiem A.

W kontekście wnioskowania statystycznego, zdarzenie A może reprezentować hipotezę, a zdarzenie B zaobserwowane dane²⁶:

$$P(H | D) = \frac{P(H) \cdot P(D | H)}{P(D)}$$

gdzie:

- $P(H)$: prawdopodobieństwo a priori — wiarygodność hipotezy przed uwzględnieniem danych,
- $P(D | H)$: wiarygodność — prawdopodobieństwo danych przy założeniu prawdziwości hipotezy,
- $P(D)$: dowód — całkowite prawdopodobieństwo danych, niezależnie od hipotezy,
- $P(H | D)$: prawdopodobieństwo a posteriori — zaktualizowane przekonanie o hipotezie po uwzględnieniu danych.

²⁵ Osvaldo Martin, *Bayesian Analysis with Python: A Practical Guide to Probabilistic Modeling* 3. wyd., (Birmingham: Packt Publishing Limited, 2024).

²⁶ Will Kurt, *Bayesian Statistics the Fun Way: Understanding Statistics and Probability with Star Wars, LEGO, and Rubber Ducks*, (San Francisco: No Starch Press, 2019).

Twierdzenie Bayesa pozwala na połączenie przekonań dotyczących danego zjawiska z zaobserwowanymi danymi i uzyskanie miary wiarygodności tych przekonań²⁷.

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

$$BF_{12} = \frac{P(D | H_1)}{P(D | H_2)}$$

Wynik ilorazu wiarygodności informuje, jak bardzo dane wspierają hipotezę H_1 w porównaniu do hipotezy H_2 . Innymi słowy, przedstawia relatywną wiarygodność danych przy założeniu prawdziwości jednej z hipotez.

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

$$O(H_1) = \frac{P(H_1)}{P(H_2)}$$

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 (ang. *posterior odds*)²⁸.

$$\frac{P(H_1 | D)}{P(H_2 | D)} = \frac{P(H_1)}{P(H_2)} \cdot \frac{P(D | H_1)}{P(D | H_2)}$$

Otrzymany wynik pokazuje, ile razy bardziej prawdopodobna jest hipoteza H_1 niż H_2 po uwzględnieniu danych D , biorąc pod uwagę zarówno wcześniejsze przekonania (iloraz prawdopodobieństw a priori), jak i zgodność danych z każdą z hipotez (czynnik Bayesa).

²⁷ Will Kurt, *Bayesian Statistics the Fun Way: Understanding Statistics and Probability with Star Wars, LEGO, and Rubber Ducks*, (San Francisco: No Starch Press, 2019).

²⁸ Will Kurt, *Bayesian Statistics the Fun Way: Understanding Statistics and Probability with Star Wars, LEGO, and Rubber Ducks*, (San Francisco: No Starch Press, 2019).

Tablica 1. Wytyczne ewaluacji ilorazu prawdopodobieństw a posteriori

Iloraz prawdopodobieństw a posteriori	Siła dowodu
1 do 3	Interesujące, ale niekonkluzywne
3 do 20	Umiarkowane dowody na korzyść H_1
20 do 150	Silne dowody na korzyść H_1
powyżej 150	Przytłaczające dowody na korzyść H_1

Źródło: Opracowanie własne na podstawie *Bayesian Statistics the Fun Way*²⁹.

Istnieje wiele innych wytycznych dotyczących ewaluacji ilorazu prawdopodobieństw a posteriori oraz czynnika 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 dane³⁰. 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

²⁹ Will Kurt, *Bayesian Statistics the Fun Way: Understanding Statistics and Probability with Star Wars, LEGO, and Rubber Ducks*, (San Francisco: No Starch Press, 2019).

³⁰ Osvaldo Martin, *Bayesian Analysis with Python: A Practical Guide to Probabilistic Modeling* 3. wyd., (Birmingham: Packt Publishing Limited, 2024).

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 określone są rozkłady a priori, reprezentujące 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, jako³¹:

$$\alpha \sim \text{Rozkład a priori}$$

$$\beta \sim \text{Rozkład a priori}$$

$$\sigma \sim \text{Rozkład a priori}$$

$$\mu = \alpha + \beta x$$

$$Y \sim \mathcal{N}(\mu, \sigma)$$

Powyższy model określa generatywny proces tworzenia danych. Parametry α , β oraz σ 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

³¹ Osvaldo Martin, *Bayesian Analysis with Python: A Practical Guide to Probabilistic Modeling* 3. wyd., (Birmingham: Packt Publishing Limited, 2024).

zmiennych objaśniających. Odchylenie standardowe można modelować jako zmienną liniowo zależną od jednej lub wielu zmiennych objaśniających. Oznacza to, że nie tylko parametr średniej (np. $\mu = \alpha + \beta x$), ale również inne parametry rozkładu warunkowego — takie jak odchylenie standardowe — mogą być modelowane jako funkcje predyktorów³².

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 (ang. *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.

2. Brak łączenia grup (ang. *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.

3. Częściowe łączenie grup (ang. *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 (ang. *shrinkage*), polegający na „przyciąganiu” estymat 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 (ang. *overfitting*). Modele hierarchiczne pozwalają na

³² Osvaldo Martin, *Bayesian Analysis with Python: A Practical Guide to Probabilistic Modeling* 3. wyd., (Birmingham: Packt Publishing Limited, 2024).

jednoczesne uchwycenie efektu globalnego (wspólnego trendu dla wszystkich grup) oraz efektów specyficznych dla poszczególnych grup³³.

Po dopasowaniu jednego lub wielu modeli konieczna jest ich ewaluacja i walidacja. Jednym z kluczowych narzędzi w ramach podejścia bayesowskiego jest weryfikacja predykcji a posteriori (ang. *posterior predictive checks*). Polega ona na ocenie stopnia dopasowania modelu do danych, poprzez generowanie nowych obserwacji na podstawie parametrów wylosowanych z rozkładów a posteriori. Innymi słowy, sprawdzana jest zgodność między danymi generowanymi przez model a danymi rzeczywistymi — analizowane są odchylenia od "prawdziwego" rozkładu danych³⁴.

W kontekście porównywania wielu modeli opisujących to samo zjawisko, szczególnie istotna jest ocena zdolności do generalizacji, czyli trafnego przewidywania nowych, niewidzianych wcześniej obserwacji. Dwa najczęściej stosowane podejścia w tym celu to:

1. Kryteria informacyjne (ang. *information criteria*)

Pojęcie to obejmuje metody aproksymujące zdolność modelu do predykcji nowych obserwacji. Typowo opierają się one na sumie: (1) miary dopasowania modelu do danych użytych do dopasowania modelu oraz (2) składnika penalizującego złożoność modelu. Popularnymi przykładami są kryterium informacyjne Akaikego (AIC) oraz bayesowskie kryterium informacyjne Schwarza (BIC). W kontekście bayesowskim często stosowane jest także WAIC (Watanabe-Akaike Information Criterion).

2. Sprawdzanie krzyżowe (ang. *cross-validation*)

Strategie polegające na empirycznym podziale danych na podzbiory oraz ocenie zdolności predykcyjnych modelu na podstawie danych nieuczestniczących w dopasowaniu parametrów. Mimo swojej intuicyjności, klasyczne warianty (np. prosty podział trening/test) bywają krytykowane za nieuzasadnione odrzucanie części informacji. Istnieją jednak bardziej zaawansowane metody, takie jak sprawdzanie k-krotne (ang. *k-fold cross-validation*), które ograniczają ten problem.

W kontekście modeli bayesowskich lub dużych zbiorów danych skuteczne metody sprawdzania krzyżowego często są poza zasięgiem mocy obliczeniowych komputerów personalnych.

³³ Richard McElreath, *Statistical rethinking: a Bayesian course with examples in R and Stan* 2. wyd. Chapman & Hall/CRC texts in statistical science series, (Boca Raton: CRC Press, 2020).

³⁴ „Prior and Posterior Predictive Checks — PyMC 5.22.0 documentation”, dostęp 15 maj 2025, https://www.pymc.io/projects/docs/en/stable/learn/core_notebooks/posterior_predictive.html.

Istnieją jednak sposoby na aproksymacje sprawdzania krzyżowego typu *leave-one-out*. Szczególnie skuteczną w praktyce metodą jest *Pareto smooth importance sampling leave-one-out cross-validation* (PSIS-LOO-CV). Umożliwia ona wiarygodną ocenę zdolności predykcyjnych modelu bez konieczności jego wielokrotnego dopasowywania.

Pomimo koncepcyjnej prostoty, metody bayesowskie często okazują się trudne do zastosowania analitycznego, zwłaszcza w przypadku modeli o złożonej strukturze i wielu parametrach. Główną trudnością jest obliczenie mianownika w twierdzeniu Bayesa, tzw. dowodu, który wymaga całkowania po całej przestrzeni parametrów. Dla wielu modeli operacja ta jest obliczeniowo nieosiągalna. W celu przybliżenia rozkładów a posteriori wykorzystuje się zatem metody numeryczne, w szczególności algorytmy z rodziny MCMC (ang. *Markov Chain Monte Carlo*). W teorii, algorytmy MCMC umożliwiają przybliżenie rozkładu a posteriori dowolnego modelu probabilistycznego, niezależnie od jego złożoności. Istnieją również podejścia analityczne, takie jak metoda siatki (ang. *grid method*) czy metoda Laplace’a (ang. *Laplace approximation*), jednak ich zastosowanie staje się szybko niepraktyczne ze względu na koszt obliczeniowy szybko rosnący wraz ze wzrostem liczby parametrów. Metody MCMC należą do grupy metod stochastycznych i pozwalają na generowanie próbek z rzeczywistego rozkładu a posteriori, nawet w przypadkach, gdy jego postać analityczna pozostaje nieznana. Ich kluczową zaletą jest efektywność w kontekście wysokowymiarowych przestrzeni parametrów – potrafią eksplorować regiony rozkładu proporcjonalnie do ich względnego prawdopodobieństwa. Oznacza to, że próbki częściej pochodzą z obszarów o wysokim prawdopodobieństwie a posteriori. Podstawą działania algorytmów MCMC są łańcuchy Markowa, czyli matematyczne obiekty opisujące sekwencje stanów, w których przejścia między stanami są definiowane przez tzw. macierz przejść. Istotną cechą łańcuchów Markowa jest to, że prawdopodobieństwo przejścia do kolejnego stanu zależy wyłącznie od stanu bieżącego, a nie od historii procesu (tzw. własność Markowa). W kontekście statystyki bayesowskiej łańcuchy Markowa stanowią formalną strukturę do analizy własności algorytmów MCMC oraz ich zbieżności do rzeczywistego rozkładu docelowego³⁵.

³⁵ Osvaldo Martin, *Bayesian Analysis with Python: A Practical Guide to Probabilistic Modeling* 3. wyd., (Birmingham: Packt Publishing Limited, 2024).

Rozdział 3. Modelowanie wpływu stresu na temperaturę i tętno na podstawie danych z urządzeń ubieralnych pielęgniarek

3.1 Eksploracyjna analiza badanego zbioru danych

Analiza 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 (*ang. electrodermal activity, EDA*) oraz temperaturę skóry. Dane były zbierane podczas pracy pielęgniarek w szpitalu w Stanach Zjednoczonych, w okresie pandemii COVID-19³⁶.

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.

³⁶ Seyedmajid Hosseini i in., „A Multi-Modal Sensor Dataset for Continuous Stress Detection of Nurses in a Hospital”, (Dryad, 17 wrzesień 2021), <https://doi.org/10.5061/DRYAD.5HQBZKH6F>.

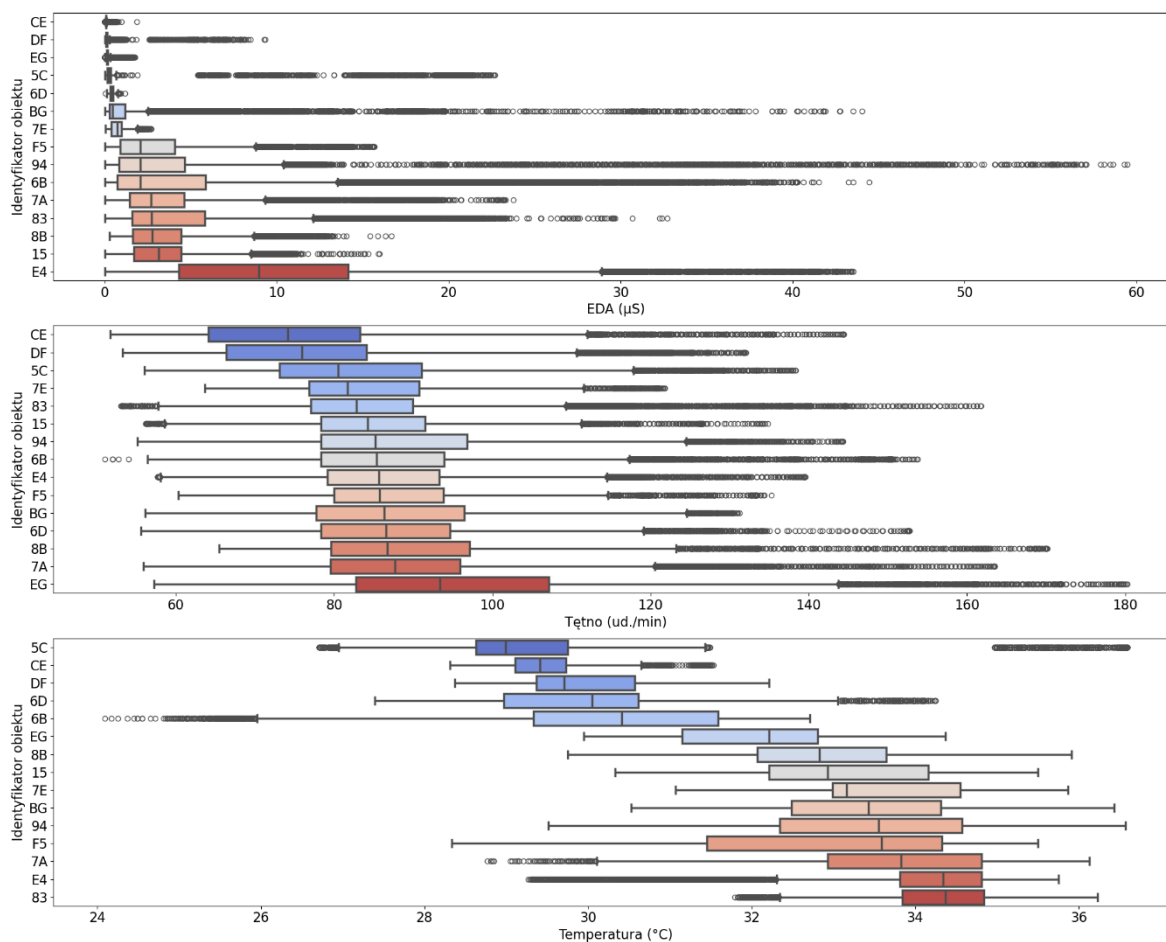
Tablica 2. Statystyki opisowe dla zmiennych ilościowych

Miara/Zmienna	Reakcja Elektrodermalna (μs)	Tętno (ud./min.)	Temperatura skóry ($^{\circ}\text{C}$)
Średnia arytmetyczna	3,51	85,76	32,25
Odchylenie standardowe	5,68	14,23	2,26
Wartość minimalna	0	51	24,09
Kwartyl pierwszy	0,22	76,68	30,21
Mediana	1,14	83,9	32,59
Kwartyl trzeci	4,06	93,13	34,27
Wartość maksymalna	59,45	180,23	36,59

Źródło: Opracowanie własne.

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

Wykres 6. Rozkłady zmiennych ilościowych w podziale na uczestników badania



Źródło: Opracowanie własne.

Wykres 6 obrazuje znaczną zmienność wartości pomiarowych pomiędzy poszczególnymi uczestnikami badania. W szczególności uczestnik E4 charakteryzuje się dużym rozproszeniem zmiennej opisującej reakcję elektrodermalną (EDA). Ponadto zaobserwowano znaczną różnorodność wartości EDA na poziomie indywidualnych uczestników.

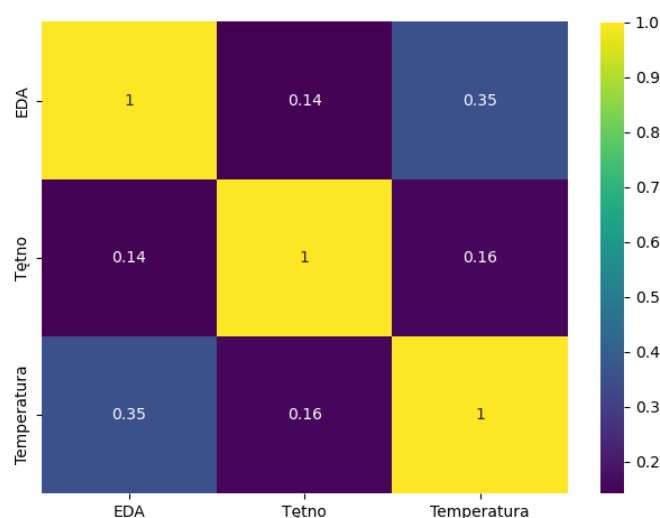
Podobne zróżnicowanie dotyczyło zmiennej opisującej tętno. Średnia arytmetyczna tętna dla uczestnika z najniższymi odczytami wynosiła 75,47 uderzeń na minutę, natomiast dla osoby z najwyższą średnią – 84,68 uderzeń na minutę.

Temperatura skóry również różnicowała badanych. Można zaobserwować dwie główne grupy uczestników 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 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 wewnątrz 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.

Wykres 7. Macierz korelacji dla zmiennych ilościowych

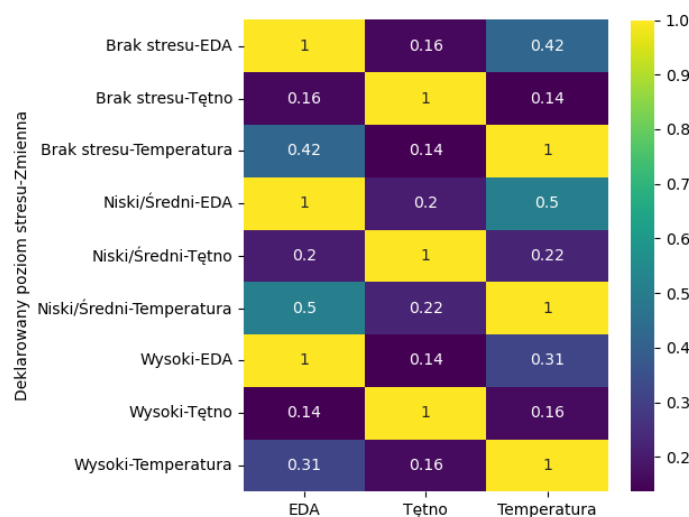


Źródło: Opracowanie własne.

Analiza korelacji wskazuje, że najsilniejszy związek występuje pomiędzy temperaturą skóry a aktywnością elektrodermalną ($r = 0,35$). Oznacza to słabą 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 ($r = 0,14$).

Ocena siły związku zmienia się po stratyfikacji uczestników ze względu na deklarowany poziom stresu.

Wykres 8. Macierz korelacji dla zmiennych ilościowych z podziałem na poziom stresu

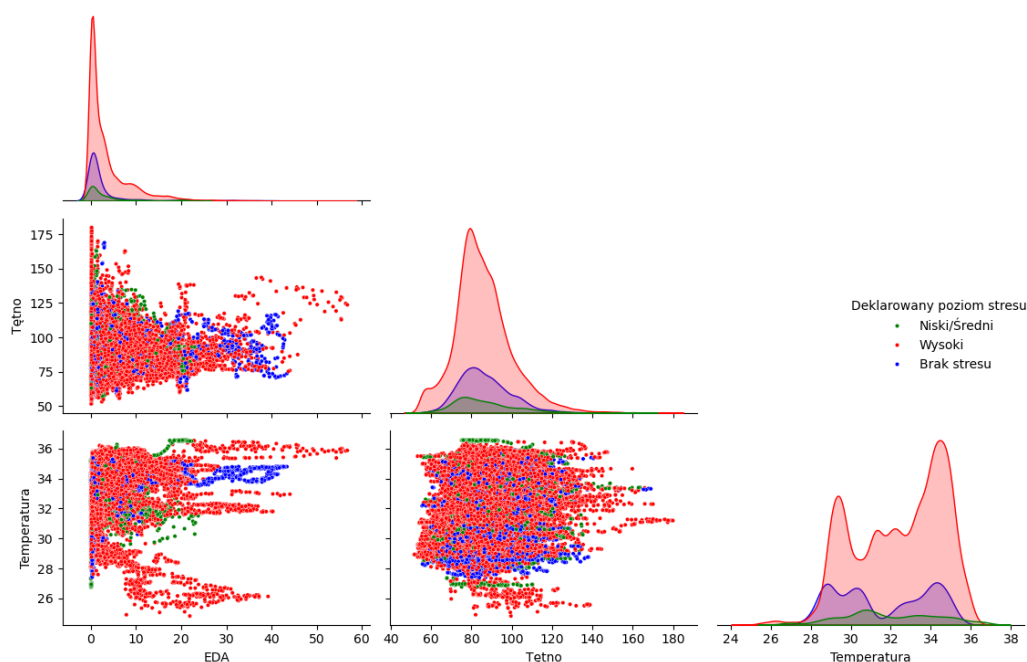


Źródło: Opracowanie własne.

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.

Wykres 9. Wykresy rozrzutu i rozkłady zmiennych ilościowych z podziałem na poziom stresu



Źródło: Opracowanie własne.

Pierwszym wnioskiem z wykresu jest zauważalna różnica w liczbie obserwacji w poszczególnych kategoriach stresu. Uczestnicy 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 wskazówek dotyczących kierunku dalszej analizy.

3.2 Wpływ deklarowanego poziomu stresu na zmienne fizjologiczne

W pierwszej kolejności uwagę skupiono na zmianach i różnicach w wartości zmiennych fizjologicznych zmierzonych wśród uczestników. Postawiono 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 modelach liniowych, dla dwóch zmiennych objaśnianych: temperatura skóry oraz tętno.

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 każdej pielęgniarce przypisywane są 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.

Model 1: Brak łączenia między grupami, liniowy efekt stresu (kodowanie porządkowe)

$$\alpha_i \sim \mathcal{N}(\bar{y}, 5)$$

$$\beta_i \sim \mathcal{N}(0, 2)$$

$$\sigma_i \sim \text{Exp}(1)$$

$$\mu_{ij} = \alpha_i + \beta_i \cdot \text{stres}_{ij}$$

$$y_{ij} \sim \mathcal{N}(\mu_{ij}, \sigma_i)$$

gdzie:

- i : indeks pielęgniarki,
- j : indeks obserwacji,
- α_i : wyraz wolny dla pielęgniarki,
- β_i : współczynnik regresji (efekt stresu) dla pielęgniarki i ,
- σ_i : odchylenie standardowe dla pielęgniarki i ,
- $\text{stres}_{ij} \in \{0, 1, 2\}$: poziom stresu (0 = Brak, 1 = Niski/Średni, 2 = Wysoki),

- y_{ij} : zmienna zależna (temperatura skóry, tętno),
- \bar{y} : średnia wartość zmiennej zależnej.

Model 2: Brak łączenia między grupami, zmienna zerojedynkowa stresu

$$\alpha_i \sim \mathcal{N}(\bar{y}, 5)$$

$$\beta_{0i} \sim \mathcal{N}(0, 2)$$

$$\beta_{1i} \sim \mathcal{N}(0, 2)$$

$$\sigma_i \sim \text{Exp}(1)$$

$$\mu_{ij} = \alpha_i + \beta_{0i} \cdot \text{Niski}\acute{\text{S}}r_{ij} + \beta_{1i} \cdot \text{Wysoki}_{ij}$$

$$y_{ij} \sim \mathcal{N}(\mu_{ij}, \sigma_i)$$

gdzie:

- i : indeks pielęgniarki,
- j : indeks obserwacji,
- α_i : wyraz wolny dla pielęgniarki i ,
- β_0 : współczynnik regresji dla poziomu stresu "Niski/Średni" dla pielęgniarki i ,
- β_1 : współczynnik regresji dla poziomu stresu "Wysoki" dla pielęgniarki i ,
- σ_i : odchylenie standardowe dla pielęgniarki i ,
- $\text{Niski}\acute{\text{S}}r_{ij} \in \{0, 1\}$: zmienna binarna wskazująca, czy poziom stresu w obserwacji j pielęgniarki i to "Niski/Średni" (1 = Tak, 0 = Nie),
- $\text{Wysoki}_{ij} \in \{0, 1\}$: zmienna binarna wskazująca, czy poziom stresu w obserwacji j pielęgniarki i to "Wysoki" (1 = Tak, 0 = Nie),
- y_{ij} : zmienna zależna (temperatura skóry, tętno),
- \bar{y} : średnia wartość zmiennej zależnej.

Model 3: Częściowe łączenie, model hierarchiczny

$$\mu_\alpha \sim \mathcal{N}(\bar{y}, 10)$$

$$\sigma_\alpha \sim \text{Exp}(1)$$

$$\mu_\beta \sim \mathcal{N}(0, 2)$$

$$\sigma_\beta \sim \text{Exp}(1.0)$$

$$\alpha_i \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha)$$

$$\beta_i \sim \mathcal{N}(\mu_\beta, \sigma_\beta)$$

$$\sigma_i \sim \text{Exp}(1.0)$$

$$\mu_{ij} = \alpha_i + \beta_i \cdot \text{stres}_{ij}$$

$$y_{ij} \sim \mathcal{N}(\mu_{ij}, \sigma_i)$$

gdzie:

- i : indeks pielęgniarki,
- j : indeks obserwacji,
- α_i : wyraz wolny dla pielęgniarki i ,
- β_i : współczynnik efektu stresu dla pielęgniarki i ,
- σ_i : odchylenie standardowe dla pielęgniarki i ,
- stres_{ij} : stres w obserwacji j dla pielęgniarki i ,
- y_{ij} : zmienna zależna (temperatura skóry, tętno),
- \bar{y} : średnia wartość zmiennej zależnej.

Model 4: Brak łączenia między grupami, zmienna wariancja

$$\alpha_i \sim \mathcal{N}(\bar{y}, 5)$$

$$\beta_{0i} \sim \mathcal{N}(0, 2)$$

$$\beta_{1i} \sim \mathcal{N}(0, 2)$$

$$\mu_{ij} = \alpha_i + \beta_{0i} \cdot \text{Niski} \hat{r}_{ij} + \beta_{1i} \cdot \text{Wysoki}_{ij}$$

$$\sigma_{\alpha,i} \sim \mathcal{N}(0, 1)$$

$$\sigma_{\beta,i} \sim \mathcal{N}(0, 1)$$

$$\log(\sigma_{ij}) = \sigma_{\alpha,i} + \sigma_{\beta,i} \cdot \text{stres}_{ij}$$

$$\sigma_{ij} = \exp(\log(\sigma_{ij}))$$

$$y_{ij} \sim \mathcal{N}(\mu_{ij}, \sigma_i)$$

gdzie:

- i : indeks pielęgniarki,
- j : indeks obserwacji,
- α_i : wyraz wolny dla pielęgniarki i ,
- β_{0i} : współczynnik regresji dla poziomu stresu "Niski/Średni" dla pielęgniarki i ,
- β_{1i} : współczynnik regresji dla poziomu stresu "Wysoki" dla pielęgniarki i ,
- σ_{ai} : punkt przecięcia dla $\log(\sigma)$ dla pielęgniarki i ,
- $\sigma_{\beta i}$: nachylenie dla $\log(\sigma)$ dla pielęgniarki i ,
- $\text{stres}_{ij} \in \{0, 1, 2\}$: poziom stresu (0 = Brak, 1 = Niski/Średni, 2 = Wysoki),
- σ_i : odchylenie standardowe dla pielęgniarki i ,
- y_{ij} : zmienna zależna (temperatura skóry, tętno),
- \bar{y} : średnia wartość zmiennej zależnej.

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

Tablica 3. Ranking testowanych specyfikacji modeli

Specyfikacja modelu	y_{ij} = temperatura skóry		y_{ij} = tętno	
	ranking	elpd_loo	ranking	elpd_loo
Model 4 (zmienna wariancja)	1	-51404,42	1	-142135,36
Model 2 (zmienna zerojedynkowa stresu)	2	-53304,68	2	-142350,85
Model 3 (model hierarchiczny)	3	-53682,22	3	-142388,34
Model 1 (zmienna stresu kodowana porządkowo)	4	-53682,48	4	-142392,7

Źródło: Opracowanie własne.

Pomimo stosunkowo niewielkich różnic w wartościach ELPD-LOO między porównywanymi modelami, szczególną uwagę zwrócono na model zakładający heteroskedastyczność, który uzyskał najwyższe wartości tej statystyki dla obu analizowanych zmiennych zależnych. Wysoka wartość ELPD-LOO sugeruje, że uwzględnienie zmiennej wariancji poprawia zdolność predykcyjną modelu. Jednocześnie analiza rozkładów a posteriori dla modeli pozbawionych struktury hierarchicznej wykazała ograniczoną interpretowalność parametrów. W związku z tym, podjęto decyzję o dalszym rozwijaniu modelu hierarchicznego, który umożliwi uwzględnienie struktury grupowej danych, a jednocześnie pozwala na modelowanie wariancji zależnej od poziomu stresu. Uzupełnieniem tej specyfikacji było zastosowanie

alternatywnej funkcji wiarygodności dla y_{ij} = tętno, która lepiej odzwierciedla empiryczny rozkład tej zmiennej.

Istotnym aspektem konstrukcji wybranego modelu była również odpowiednia specyfikacja rozkładów a priori. Zastosowano rozkłady słabo informatywne, które zapewniają umiarkowane ograniczenie przestrzeni parametrów bez narzucania silnych założeń. Przykładowo, dla parametru μ_α przyjęto rozkład normalny z wartością oczekiwaną równą średniej empirycznej zmiennej zależnej oraz odchyleniem standardowym 0,5, co umożliwia rozsądną elastyczność eksploracji przestrzeni parametrów a posteriori i jednocześnie sprzyja stabilności próbkowania. Rozkłady wykładnicze dla parametrów σ_α , μ_{β_0} , μ_{β_1} i v zapewniają dodatnie wartości, eliminując ryzyko generowania nieprawidłowych (ujemnych) wartości odchyłeń standardowych i stopni swobody.

Wybór parametrów rozkładów a priori był wynikiem iteracyjnego procesu. Początkowe eksperymenty z szerszymi, mniej informatywnymi rozkładami prowadziły do problemów z mieszalnością łańcuchów MCMC oraz wzrostu liczby rozbieżności. W odpowiedzi na te trudności zdecydowano się na bardziej informatywne, ale nadal względnie elastyczne rozkłady, które zapewniły lepszą stabilność estymacji. Natomiast w kontekście porównań modeli (w tym modeli bez struktury hierarchicznej), w obliczu braku silnej wiedzy dziedzinowej, przyjęto wspólne, mało informatywne rozkłady a priori dla wszystkich wariantów, aby zapewnić równe warunki oceny ich zdolności predykcyjnej.

Chociaż ostatecznie zaproponowano nową wersję modelu, nie oznacza to jednoznacznego odrzucenia wcześniejszych podejść. Model ten łączy zalety poprzednich rozwiązań: wykorzystuje strukturę hierarchiczną znaną z modelu 3, wykorzystuje założenie zmiennej wariancji (model 4), a także korzysta z kodowania zerojedynkowego zmiennej stresu (model 2). Dzięki analizie wyników porównań oraz ograniczeń poszczególnych wariantów, piąty model stanowi syntezę najlepszych praktyk wyłonionych w toku iteracyjnego procesu modelowania.

Model 5: Częściowe łączenie, zmienna wariancja

$$\mu_\alpha \sim \mathcal{N}(\bar{y}, 0.5)$$

$$\sigma_\alpha \sim \text{Exp}(1)$$

$$\mu_{\beta_0}, \mu_{\beta_1} \sim \mathcal{N}(0, 1)$$

$$\sigma_{\beta_0}, \sigma_{\beta_1} \sim \text{Exp}(1.0)$$

$$\alpha_i \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha)$$

$$\beta_{0i} \sim \mathcal{N}(\mu_{\beta_0}, \sigma_{\beta_0})$$

$$\beta_{1i} \sim \mathcal{N}(\mu_{\beta_1}, \sigma_{\beta_1})$$

$$\alpha_\sigma \sim \mathcal{N}(0,1)$$

$$\beta_{0\sigma}, \beta_{1\sigma} \sim \mathcal{N}(0,1)$$

$$\mu_{ij} = \alpha_i + \beta_{0i} \cdot \text{Niski}/\acute{\text{Sredni}}_{ij} + \beta_{1i} \cdot \text{Wysoki}_{ij}$$

$$\log(\sigma_{ij}) = \alpha_\sigma + \beta_{0\sigma} \cdot \text{Niski}/\acute{\text{Sredni}}_{ij} + \beta_{1\sigma} \cdot \text{Wysoki}_{ij}$$

$$\sigma_{ij} = \exp(\log(\sigma_{ij}))$$

$$v \sim \text{Exp}(1), \quad \text{jeżeli } y = \text{Tętno}$$

$$y_{ij} \sim \text{StudentT}(\mu_{ij}, \sigma_{ij}, v), \quad \text{jeżeli } y = \text{Tętno}$$

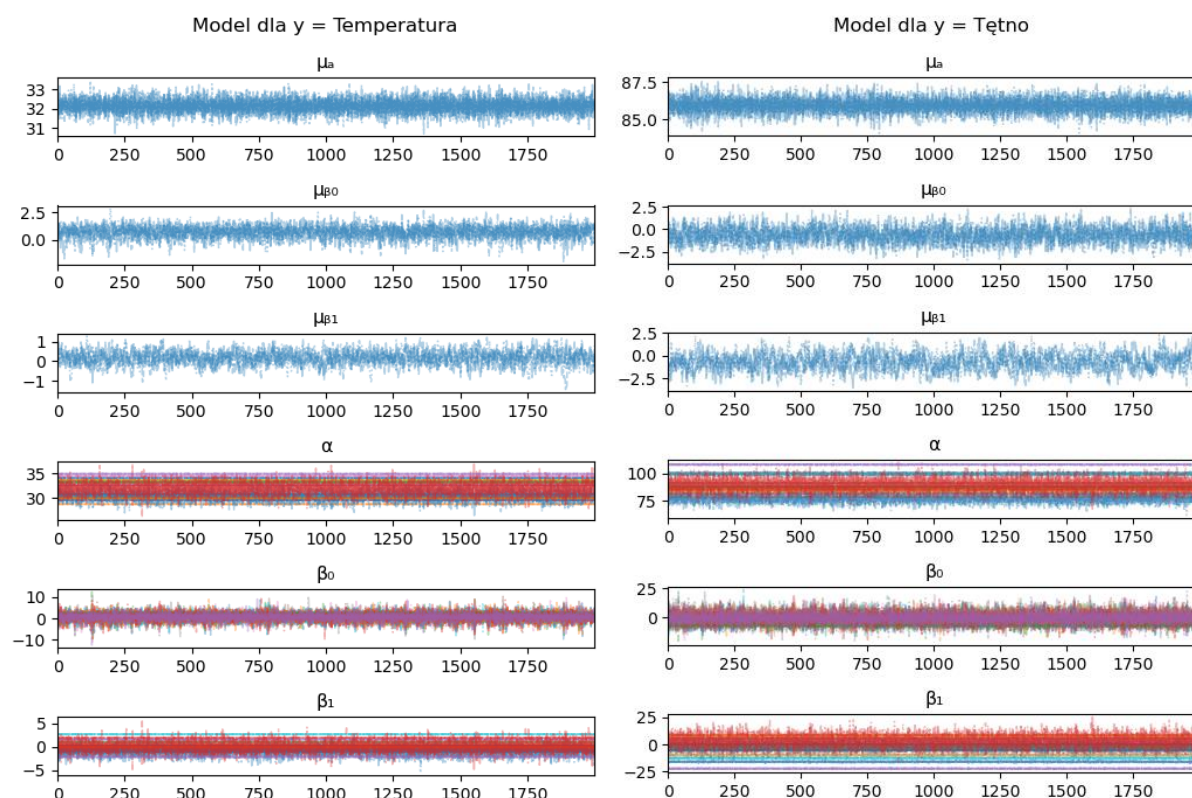
$$y_{ij} \sim \mathcal{N}(\mu_{ij}, \sigma_{ij}), \quad \text{w przeciwnym razie}$$

gdzie:

- i : indeks pielęgniarki,
- j : indeks obserwacji,
- α_i : wyraz wolny dla pielęgniarki i ,
- β_{0i} : współczynnik regresji dla poziomu stresu "Niski/Średni" dla pielęgniarki i ,
- β_{1i} : współczynnik regresji dla poziomu stresu "Wysoki" dla pielęgniarki i ,
- σ_α : punkt przecięcia dla $\log(\text{sigma})$,
- σ_{β_0} : nachylenie dla $\log(\text{sigma})$ dla poziomu stresu "Niski/Średni",
- σ_{β_1} : nachylenie dla $\log(\text{sigma})$ dla poziomu stresu "Wysoki",
- $\text{Niski}\acute{\text{S}}r_{ij} \in \{0, 1\}$: zmienna binarna wskazująca, czy poziom stresu w obserwacji j pielęgniarki i to "Niski/Średni" (1 = Tak, 0 = Nie),
- $\text{Wysoki}_{ij} \in \{0, 1\}$: zmienna binarna wskazująca, czy poziom stresu w obserwacji j pielęgniarki i to "Wysoki" (1 = Tak, 0 = Nie),
- σ_{ij} : odchylenie standardowe dla pielęgniarki i obserwacji j ,
- y_{ij} : zmienna zależna (temperatura skóry, tętno).

Analizę wyników modelowania rozpoczęto od weryfikacji procesu próbkowania MCMC, w celu ewaluacji wiarygodności uzyskanych rozkładów a posteriori. W tym celu wygenerowano wykresy trajektorii łańcuchów MCMC dla wybranych parametrów modelu, umożliwiające ocenę mieszania łańcuchów oraz potwierdzenie ich konwergencji.

Wykres 10. Wykresy trajektorii łańcuchów MCMC dla wybranych parametrów

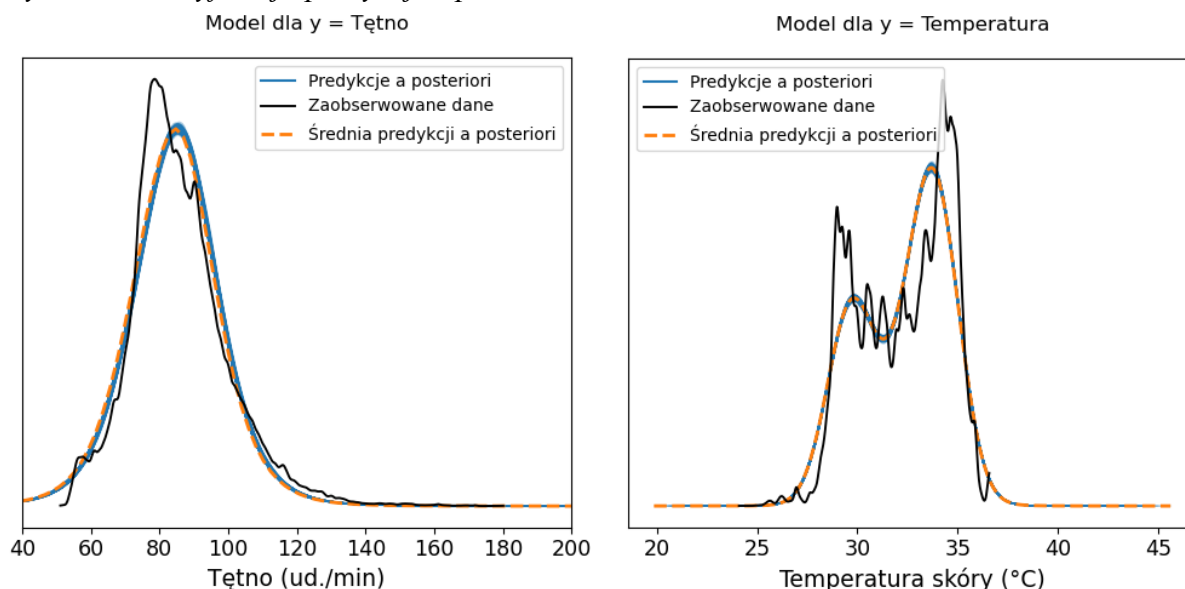


Źródło: Opracowanie własne.

Analiza wykresów trajektorii nie wykazała problemów związanych z niewystarczającym mieszaniem łańcuchów ani z dywergencjami. Poszczególne łańcuchy wykazują stabilność i dobrą mieszalność, co potwierdza, że proces próbkowania osiągnął stan stacjonarny. Na tej podstawie można uznać, że uzyskane rozkłady a posteriori są wiarygodne i nadają się do dalszej analizy.

Przeprowadzono weryfikację predykcji a posteriori w celu oceny dopasowania modelu do danych.

Wykres 11. Weryfikacja predykcji a posteriori

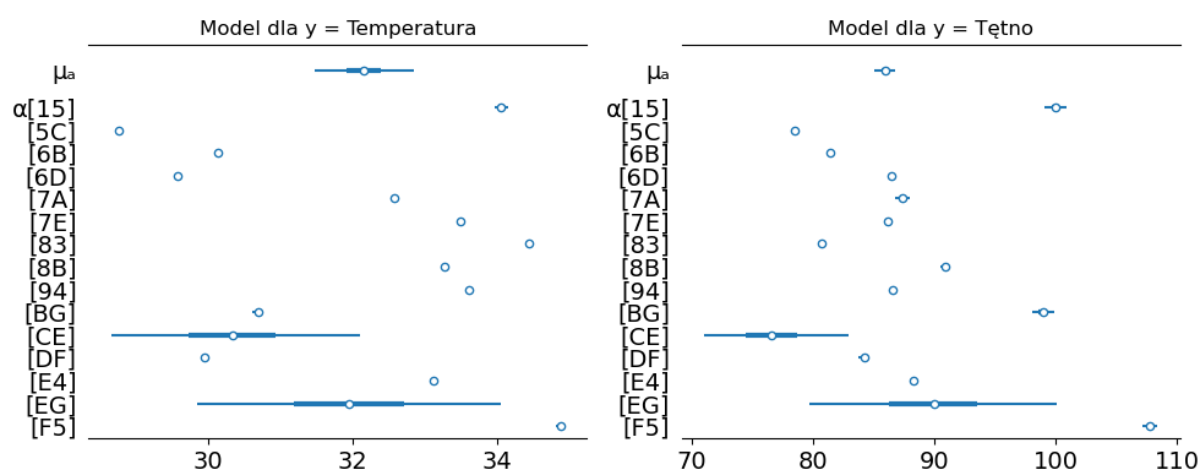


Źródło: Opracowanie własne.

Rozkłady uzyskane w wyniku predykcji a posteriori są zbliżone do rozkładów empirycznych. Sugeruje to, że model dobrze odwzorowuje strukturę danych i posiada zdolność do generowania realistycznych obserwacji. Zbieżność między rozkładami obserwowanymi i symulowanymi stanowi argument na rzecz poprawności przyjętej specyfikacji modelu.

W dalszym etapie zbadano rozkłady a posteriori dla parametrów globalnych μ_α , a także parametrów indywidualnych α , które odpowiadają częściowo łączonym estymatom wartości zmiennej zależnej w warunkach braku odczuwanego stresu przez uczestników. Analiza tych parametrów pozwala zrozumieć, jak kształtują się wartości bazowe zmiennej zależnej oraz jakie są indywidualne odchylenia od tej średniej w stanie minimalnego stresu.

Wykres 12. Rozkłady a posteriori parametrów μ_α oraz α



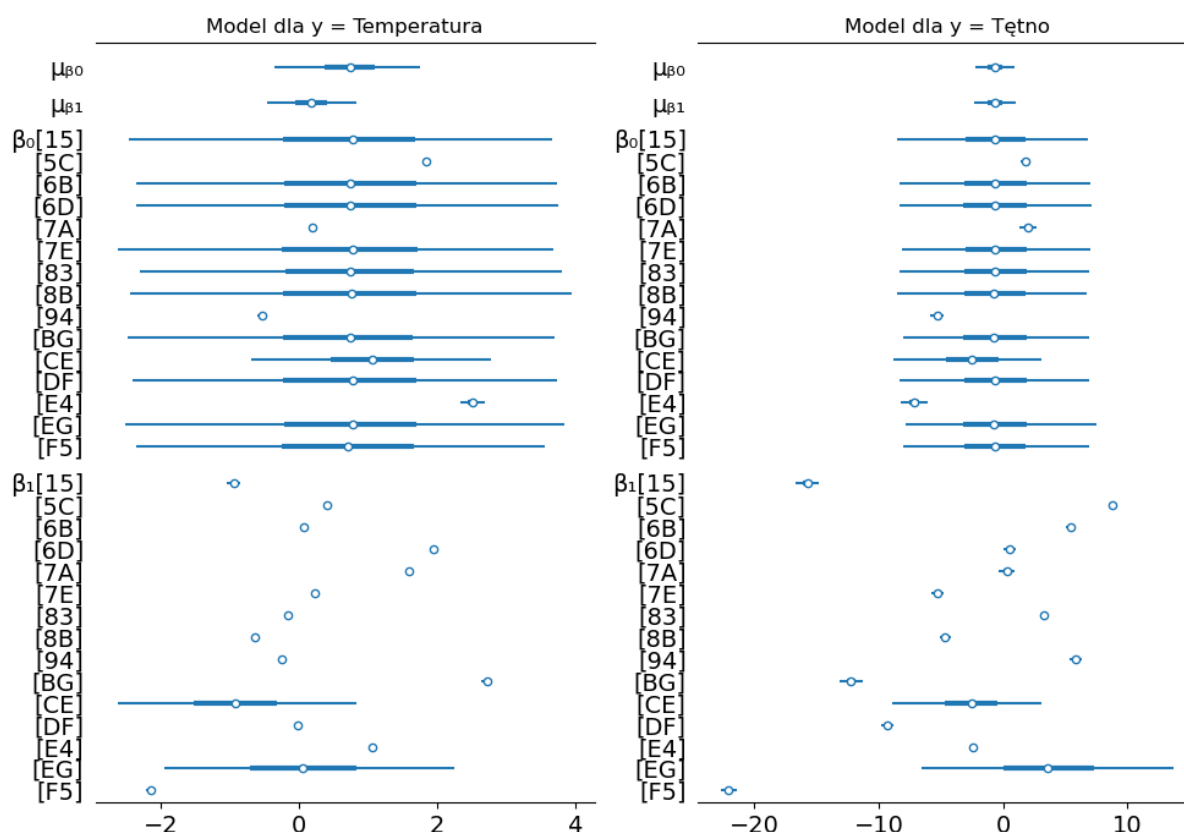
Źródło: Opracowanie własne.

Parametr globalny μ_α , reprezentujący rozkład a posteriori średniego poziomu zmiennej zależnej w sytuacjach braku stresu uczestników, wynosił 32,16°C (94% HDI: 31,48 – 32,84) dla temperatury ciała oraz 85,93 ud./min (94% HDI: 85,02 – 86,81) dla tętna.

Wśród parametrów indywidualnych α , obserwuje się znaczące zróżnicowanie między uczestnikami, zarówno dla temperatury, jak i tętna. Przykładowo, uczestnik oznaczony jako F5 cechował się najwyższym poziomem temperatury spoczynkowej (średnia rozkładu a posteriori: 34,87°C) oraz najwyższym tętnem (107,8 ud./min). Z kolei najniższe wartości odnotowano dla uczestników 5C (temperatura: 28,77°C, tętno: 78,52 ud./min) oraz CE (temperatura: 30,35°C, tętno: 76,59 ud./min), przy czym w przypadku uczestników CE i EG odnotowano również szerokie przedziały wiarygodności, co sugeruje większą niepewność estymacji tych parametrów, spowodowaną małą liczbą obserwacji, w których nie odczuwali stresu.

Następnie analizowano rozkłady a posteriori parametrów opisujących zmianę poziomu zmiennej zależnej w reakcji na stres. Parametry globalne μ_{β_0} oraz μ_{β_1} reprezentują średni efekt poziomu stresu na temperaturę ciała oraz tętno, przy czym μ_{β_0} opisuje oczekiwaną zmianę tych zmiennych w sytuacji niskiego lub średniego stresu, natomiast μ_{β_1} odnosi się do sytuacji wysokiego stresu. Z kolei parametry β_0 oraz β_1 odzwierciedlają indywidualne efekty stresu dla poszczególnych uczestników badania.

Wykres 13. Rozkłady a posteriori parametrów μ_{β_0} , μ_{β_1} , β_0 oraz β_1



Źródło: Opracowanie własne.

W przypadku temperatury ciała, średni efekt niskiego/średniego stresu (μ_{β_0}) wynosił $0,71^{\circ}\text{C}$ (94% HDI: $-0,36 - 1,75$), natomiast dla wysokiego stresu (μ_{β_1}) $0,17^{\circ}\text{C}$ (94% HDI: $-0,46 - 0,83$). Choć średnie wartości rozkładów a posteriori są dodatnie, przedziały wiarygodności obejmują zero, co sugeruje brak istotnego efektu stresu na temperaturę skóry na poziomie globalnym.

W modelu tętna, średni wpływ niskiego/średniego stresu oszacowano na $-0,63$ ud./min (94% HDI: $-2,29 - 0,90$), a dla stresu wysokiego na $-0,66$ ud./min (94% HDI: $-2,36 - 1,03$). Ponownie, przedziały wiarygodności obejmują zero, co wskazuje na brak silnych dowodów na istnienie wpływu stresu na tętno na poziomie globalnym.

Obserwowane są z kolei znaczne różnice między pielęgniarkami, w indywidualnej odpowiedzi na stres. Na przykład uczestnik oznaczony jako E4 wykazywał istotny wzrost temperatury ciała w warunkach niskiego/średniego stresu (średni wzrost o $2,5^{\circ}\text{C}$, 94% HDI: $2,33 - 2,68$), przy jednoczesnym spadku tętna o $7,12$ ud./min (94% HDI: $-8,24 - -6,07$). Uczestnik F5, mimo najwyższych wartości spoczynkowych, reagował odwrotnie – wykazywał wyraźny spadek tętna w warunkach wysokiego stresu ($-22,07$ ud./min, 94% HDI: $-22,71 - -21,44$) oraz obniżenie temperatury ciała ($-2,14^{\circ}\text{C}$, 94% HDI: $-2,20 - -2,08$). Ostatecznie, dla temperatury

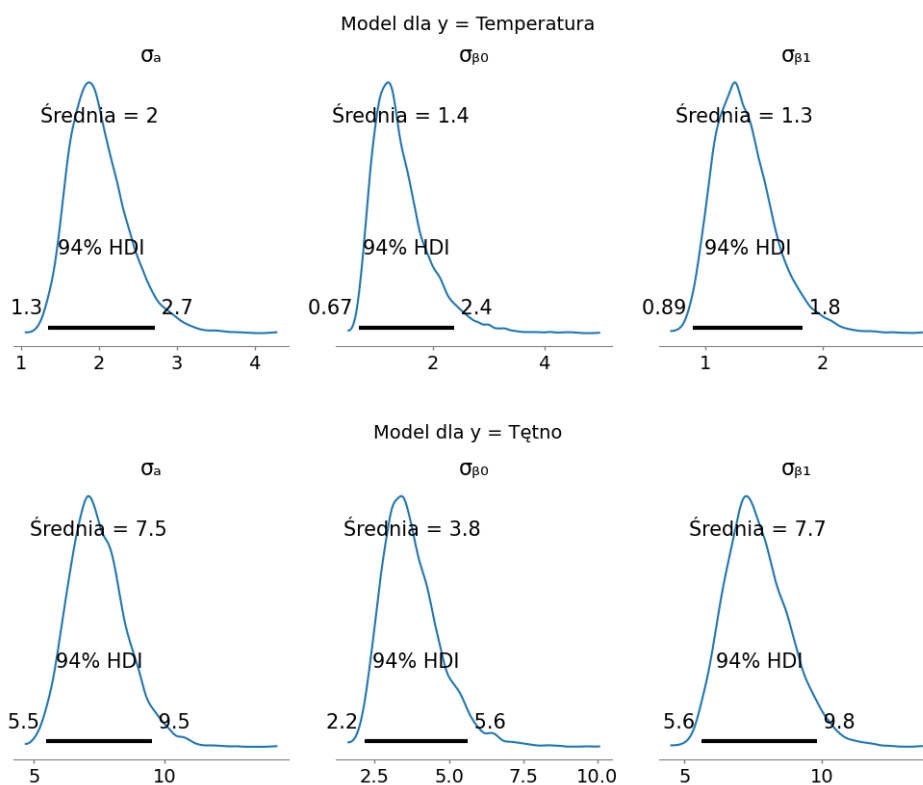
ciała, w warunkach niskiego/średniego stresu (β_0), wiarygodnie dodatni efekt (94% HDI > 0) zaobserwowano u trzech z piętnastu uczestników, natomiast wiarygodnie ujemny efekt wystąpił u jednego z uczestników. W warunkach wysokiego stresu (β_1), wiarygodnie dodatni efekt temperatury odnotowano u siedmiu uczestników, zaś ujemny efekt u pięciu.

W przypadku tętna, dla β_0 (niski/średni stres), zarówno dla efektu wiarygodnie dodatniego, jak i ujemnego odnotowano dwóch uczestników. Dla β_1 (wysoki stres), wiarygodnie dodatni efekt zaobserwowano u pięciu uczestników, natomiast ujemny u siedmiu uczestników.

Warto zauważyć, że dla części uczestników rozkłady a posteriori parametrów β_0 cechowały się dużą niepewnością i bardzo szerokimi przedziałami wiarygodności. Wynika to z faktu, że w określonych warunkach stresowych nie zgromadzono żadnych obserwacji, co uniemożliwiało precyzyjną estymację na poziomie indywidualnego efektu.

Przeanalizowano rozkłady a posteriori parametrów modelujących heteroskedastyczność: σ_α , σ_{β_0} oraz σ_{β_1} . Parametry te opisują odpowiednio wyraz wolny oraz współczynniki regresji w modelu $\log(\sigma_{ij})$. Na wykresie 14 zaprezentowano rozkłady a posteriori, łącznie z 94-procentowymi przedziałami wiarygodności oraz średnią wartością każdego rozkładu.

Wykres 14. Rozkłady a posteriori parametrów σ_α , σ_{β_0} , σ_{β_1}



Źródło: Opracowanie własne.

W przypadku modelu dla temperatury skóry, parametry σ_α , σ_{β_0} , σ_{β_1} przyjmują dodatnie wartości, co wskazuje na zróżnicowanie wariancji reszt modelu w zależności od poziomu stresu. Zróżnicowanie obserwowanych wartości temperatury wzrasta wraz z rosnącym poziomem stresu. Analogiczny wzorzec zaobserwowano w modelu dla częstości akcji serca, przy czym wartości parametrów odpowiadających heteroskedastyczności są wyraźnie wyższe, co sugeruje jeszcze silniejsze zróżnicowanie wariancji w zależności od poziomu stresu.

Zakończenie

W pracy podjęto się eksploracji potencjału idei *Quantified Self*, technologii ubieralnych oraz zrozumienia, w jaki sposób dane gromadzone przez urządzenia ubieralne mogą wspierać analizę fizjologicznych reakcji na stres w rzeczywistych warunkach pracy. W ostatnich latach obserwowano znaczący wzrost popularności urządzeń i technologii ubieralnych, które stały się narzędziem polepszania zdrowia, samopoczucia, ale również narzędziem pozwalającym osiągać lepsze wyniki sportowe, lub monitorować zdrowie pacjentów. Celem pracy było określenie wpływu poziomu stresu na reakcje fizjologiczne organizmu przy wykorzystaniu modelowania bayesowskiego oraz danych zgromadzonych przez urządzenia ubieralne. W rozdziale trzecim przeprowadzono analizę danych zebranych od personelu pielęgniarskiego pracującego w okresie pandemii w środowisku szpitalnym — zawodzie o wysokim poziomie obciążenia fizycznego i psychicznego. Wstępna eksploracja danych wskazała znaczące indywidualne zróżnicowanie rozkładów zmiennych fizjologicznych wśród uczestników badania. Badano również korelację pomiędzy tymi zmiennymi, które były na ogół dodatnie, choć przeważnie słabe do umiarkowanych. Warto również podkreślić, że dane zawierały znacznie więcej obserwacji z okresów wysokiego stresu niż umiarkowanego czy niskiego, co obrazuje wymagającą charakterystykę pracy personelu szpitalnego.

Pytania badawcze dotyczyły tego, czy i w jaki sposób zmienne fizjologiczne, takie jak tętno czy temperatura skóry, zmieniają się w zależności od deklarowanego poziomu stresu. Na podstawie porównań różnych specyfikacji modeli, wybrano do zastosowania hierarchiczny model liniowy zakładający heteroskedastyczność wariancji, który umożliwił zarówno analizę efektów globalnych, jak i indywidualnych różnic pomiędzy uczestnikami badania. Pomimo tego, że w obu modelach globalny efekt stresu nie różnił się znacznie od zera, to parametry na poziomie indywidualnym pokazują zróżnicowaną reakcję na stres. U części pielęgniarek wysoki poziom stresu wiązał się ze wzrostem wartości zmiennych fizjologicznych (przyspieszone tętno lub wzrost temperatury ciała), podczas gdy u innych obserwowano spadek lub brak wyraźnych zmian. Tego rodzaju indywidualna zmienność sugeruje, że reakcje na stres są silnie uzależnione od osobistych uwarunkowań biologicznych ludzi.

Dostępność, niski koszt oraz możliwość ciągłego rejestrowania parametrów fizjologicznych przez urządzenia ubieralne otwierają nowe możliwości ich zastosowań, między innymi w sporcie, służbie zdrowia, ale również w badaniach naukowych nad zachowaniami i stanem zdrowia ludzi w warunkach codziennych.

Bibliografia

- Carbone, Leandro, Matias Sampietro, Agustin Cicognini, Manuel García-Sillero, i Salvador Vargas-Molina. „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>.
- Catapult Support. „What Is Player Load?”, 31 styczeń 2024. <https://support.catapultsports.com/hc/en-us/articles/360000510795-What-is-Player-Load>.
- Choe, Eun Kyoung, Nicole B. Lee, Bongshin Lee, Wanda Pratt, i Julie A. Kientz. „Understanding Quantified-Selfers’ Practices in Collecting and Exploring Personal Data”. W *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1143–52. Toronto: ACM, 2014. <https://doi.org/10.1145/2556288.2557372>.
- Clayton, Aubrey. *Bernoulli’s Fallacy: Statistical Illogic and the Crisis of Modern Science*. New York: Columbia University Press, 2021.
- Ebbelaar, Dave. „Exploring the Possibilities of Context-Aware Applications for Strength Training”. Vrije Universiteit, 2019. <https://github.com/daveebelaar/tracking-barbell-exercises/tree/master>.
- Gabbett, Tim J. „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>.
- Henriksen, André, Ashenafi Zebene Woldaregay, Miroslav Muzny, Gunnar Hartvigsen, Laila Arnesdatter Hopstock, i Sameline Grimsgaard. „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>.
- Hoogendoorn, Mark, 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>.
- Hosseini, Seyedmajid, Satya Katragadda, Ravi Teja Bhupatiraju, Ziad Ashkar, Christoph Borst, Kenneth Cochran, i Raju Gottumukkala. „A Multi-Modal Sensor Dataset for Continuous Stress Detection of Nurses in a Hospital”. Dryad, 17 wrzesień 2021. <https://doi.org/10.5061/DRYAD.5HQBZKH6F>.
- Junata, Melisa, i Raymond Kai-Yu Tong. „Chapter 1 - Wearable Technology in Medicine and Health Care: Introduction”. W *Wearable Technology in Medicine and Health Care*, zredagowane przez Raymond Kai-Yu Tong, 1–5. Academic Press, 2018. <https://doi.org/10.1016/B978-0-12-811810-8.00001-4>.
- Kennedy, Brett. *Outlier Detection in Python*. 1. wyd. New York: Manning Publications Co. LLC, 2025.
- Kurt, Will. *Bayesian Statistics the Fun Way: Understanding Statistics and Probability with Star Wars, LEGO, and Rubber Ducks*. San Francisco: No Starch Press, 2019.
- Lupton, Deborah. *The Quantified Self: A Sociology of Self-Tracking*. Cambridge: Polity, 2016.
- Mann, Steve. „Mediated Reality”. *Linux Journal* 1999, nr 59es (1 marzec 1999): 5-es.
- . „Wearable computing as means for personal empowerment”. W *Proc. 3rd Int. Conf. on Wearable Computing (ICWC)*, 51–59, 1998. 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.
- Martin, Osvaldo. *Bayesian Analysis with Python: A Practical Guide to Probabilistic Modeling*. 3. wyd. Birmingham: Packt Publishing Limited, 2024.

- McElreath, Richard. *Statistical rethinking: a Bayesian course with examples in R and Stan*. 2. wyd. Chapman & Hall/CRC texts in statistical science series. Boca Raton: CRC Press, 2020.
- Myatt, Glenn J., i Wayne P. Johnson. *Making Sense of Data I: A Practical Guide to Exploratory Data Analysis and Data Mining*. 2. wyd. Safari Tech Books Online. Hoboken: Wiley, 2014.
- Nielsen. „Hacking Health: How Consumers Use Smartphones and Wearable Tech to Track Their Health”. Dostęp 27 listopad 2024.
<https://www.nielsen.com/insights/2014/hacking-health-how-consumers-use-smartphones-and-wearable-tech-to-track-their-health/>.
- „Prior and Posterior Predictive Checks — PyMC 5.22.0 documentation”. Dostęp 15 maj 2025.
https://www.pymc.io/projects/docs/en/stable/learn/core_notebooks/posterior_predictive.html.
- 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>.
- Ranga Suri, N. N. R. *Outlier Detection: A Data Mining Perspective*. Intelligent Systems Reference Library, v. 155. Cham: Springer, 2019.
- Slevin, Patrick, i Brian Caulfield. „Chapter 13 - Patient-Generated Health Data: Looking Toward Future Health Care”. W *Wearable Technology in Medicine and Health Care*, zredagowane przez Raymond Kai-Yu Tong, 261–73. Academic Press, 2018.
<https://doi.org/10.1016/B978-0-12-811810-8.00013-0>.
- Tukey, John Wilder. *Exploratory Data Analysis*. Addison-Wesley Series in Behavioral Science. Reading: Addison-Wesley publ, 1977.
- Wearable technologies and sports analytics, University of Michigan*. T. Sports Performance Analytics. Coursera. Michigan: Coursera, 2021.
- Wolf, Gary. „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/>.
- Worsey, Matthew T. O., Hugo G. Espinosa, Jonathan B. Shepherd, i David V. Thiel. „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>.

Spis wykresów

Wykres 1. Liczba urządzeń ubieralnych wprowadzanych na rynek w latach 2011-2017 (I połowa).....	8
Wykres 2. Procent urządzeń ubieralnych wykorzystujących poszczególne sensory w latach 2011-2017.....	9
Wykres 3. Liczba publikacji na temat urządzeń ubieralnych w kontekście opieki zdrowotnej i medycyny na PubMed (2005–2024)	13
Wykres 4. Rozkład syntetycznych zmiennych x i y	17
Wykres 5. Związek zmiennych x i y	18
Wykres 6. Rozkłady zmiennych ilościowych w podziale na uczestników badania	29
Wykres 7. Macierz korelacji dla zmiennych ilościowych	30
Wykres 8. Macierz korelacji dla zmiennych ilościowych z podziałem na poziom stresu	31
Wykres 9. Wykresy rozrzutu i rozkłady zmiennych ilościowych z podziałem na poziom stresu	32
Wykres 10. Wykresy trajektorii łańcuchów MCMC dla wybranych parametrów	39
Wykres 11. Weryfikacja predykcji a posteriori	40
Wykres 12. Rozkłady a posteriori parametrów μ_α oraz α	40
Wykres 13. Rozkłady a posteriori parametrów $\mu_{\beta_0}, \mu_{\beta_1}, \beta_0$ oraz β_1	42
Wykres 14. Rozkłady a posteriori parametrów $\sigma_\alpha, \sigma_{\beta_0}, \sigma_{\beta_1}$	43

Spis tablic

Tablica 1. Wytyczne ewaluacji ilorazu prawdopodobieństw a posteriori	23
Tablica 2. Statystyki opisowe dla zmiennych ilościowych	29
Tablica 3. Ranking testowanych specyfikacji modeli	36