

Akademia Wychowania Fizycznego
im. Jerzego Kukuczki w Katowicach

mgr Tomasz Wilczewski

Wskaźniki gry determinujące wynik rywalizacji sportowej
w świetle analizy rozgrywek na najwyższym poziomie
mistrzostwa w koszykówce

Autoreferat rozprawy na stopień doktora nauk o kulturze fizycznej

Promotor

dr hab. Kazimierz Mikołajec prof. AWF Katowice

KATOWICE 2023

Spis treści

Spis treści	2
WSTĘP	3
1. PODSTAWY TEORETYCZNE PROBLEMU BADAWCZEGO	4
2. CEL PRACY I PYTANIA BADAWCZE	5
3. MATERIAŁ I METODY BADAŃ.....	6
3.2 Techniki i narzędzia badawcze	8
3.2.1 Autorski arkusz obserwacji przebiegu gry	8
3.2.2 Analiza rzetelności – porównanie obserwatorów	9
3.3 Modułowe narzędzie do rejestracji i analizy działań ofensywnych.....	13
3.3.1 Pierwszy moduł analiz – dane podstawowe	13
3.3.2 Drugi moduł analiz – przedziały czasowe	14
3.3.3 Trzeci moduł analiz – ciągi rzędowe	16
3.4 Narzędzia analizy statystycznej	19
4. WYNIKI BADAŃ.....	22
4.1 Redukcja liczby zmiennych – wybór zmiennych do badań	22
4.2 Analiza zmiennych diagnostycznych – korelacje rang Spearmana	24
4.3 Klasyfikacja zespołów do grup – analiza dyskryminacyjna	27
4.4 Ranking teoretyczny i rzeczywisty – porządkowanie liniowe Hellwiga	30
4.5 Jednorodne podzbiory mistrzostw – analiza taksonomiczna	32
4.6 Prognozy zwycięstwa i porażki – regresja logistyczna	33
5. WNIOSKI.....	34
BIBLIOGRAFIA.....	36

WSTĘP

Koszykówka jest współcześnie dyscypliną sportu o niezwykle rozbudowanych systemach taktycznych. Z tego względu wzrosło znaczenie współdziałania podczas gry w oparciu o wszechstronne wyszkolenie techniczne poszczególnych zawodników¹. Znacznie zwiększyła się też liczba elementów techniki ataku i obrony, ponieważ poszukuje się coraz skuteczniejszych możliwości działania w czasie rywalizacji².

Analiza gry jest jedną z najbardziej istotnych form diagnostycznych stosowanych w grach zespołowych, a jej rezultaty mają decydujący wpływ na planowanie, realizację cykli treningowych oraz osiągnięcie optymalnych wyników w rywalizacji sportowej³. Zdecydowana większość przeprowadzonych do tej pory analiz została dokonana w oparciu o dane zarejestrowane w lidze NBA⁴ lub Eurolidze. Grają w nich najlepsi zawodnicy, pracują najwybitniejsi trenerzy i to tam, najwcześniej pojawiają się rewolucyjne koncepcje treningowe i meczowe. Liczba rejestrowanych elementów gry w ataku i obronie jest niezwykle długa i publikowana na stronach internetowych, dzięki czemu jest dostępna dla szerokiego grona zainteresowanych z całego świata.

Odkrywanie przestrzeni niedostępnych dla przeciętnego sympatyka koszykówki w jej najmniejszych składowych to ciąg fascynujących przeżyć i doświadczeń. Wydawać by się mogło, że liczba możliwości taktycznych, wzorców zachowań i indywidualnych działań zawodników jest nieskończona. Korelując ze sobą poszczególne struktury coraz wyższego rzędu, uzyskuje się prawdopodobieństwo dotarcia do źródła, do doskonałości. Prognozowanie, korelacje, tendencje i istotności różnych zmiennych diagnostycznych to namiętności wielu badaczy i teoretyków na całym świecie. Prezentowana praca podejmuje próbę odnalezienia rozwiązań doskonałych nie tylko w obszarze teoretyczno-badawczym, ale również przyjmuje kierunek realizacji swoich wyników do świata codziennych praktyk w uczestnictwie sportowym dzieci, młodzieży oraz sportu zawodowego.

¹ A. Szwarc, I. Lekner, P. Lipińska, *The efficiency of action of the top level basketball players with regard to the player position*. AWF, Wrocław 2013, nr 42, s. 44 – 61.

² J. Dembiński, *Kwantyfikacja potencjału działania zespołowego a potencjał indywidualny na przykładzie koszykówki cz. I. Założenia teoretyczne*, „Edukacja Wychowanie Sport”, 2021, nr 2.

³ T. Huciński, M. Czerlonko, *Ocena efektywności walki sportowej zespołu w koszykówce*, Roczniki naukowe AWF, Gdańsk 2001, s. 27 – 37.

⁴ K. Mikołajec, A. Maszczyk, T. Zając, *Game indicators Determining Sports Performance in the NBA*, „Journal of Human Kinetics”, 2013, vol. 37, s. 145 – 151.

Zdecydowana większość publikacji z dziedziny koszykówki skupia się wokół zagadnień skuteczności rzutowej wynikającej z rodzaju rzutu, techniki jego wykonania, odległości od kosza oraz cech somatycznych zawodników. Sporo uwagi w ostatnich latach poświęca się też programom treningowym skuteczności rzutowej⁵ i testowaniu innowacyjnych metod w celu jej poprawy. Zbadano na przykład efekty treningu wizualizacji⁶ czy też zastosowanie nietradycyjnych pomocy i środków treningowych, tj. specjalnych piłek i usztywniaczy ręki nierzucającej⁷ oraz zmniejszonych średnicy obręczy do sprawdzania skuteczności rzutów wolnych⁸. W obszarze zainteresowań naukowców znalazł się również wpływ wprowadzenia piłek do koszykówki o różnej masie i średnicy na skuteczność rzutową podczas zawodów na różnym poziomie sportowym⁹.

Konkludując, zasadniczym, pierwszym założeniem niniejszej pracy było stworzenie nowego narzędzia do wielowymiarowej analizy gry w aspektach dotąd nie odkrytych, nie branych pod uwagę w konwencjonalnym procesie analizy. Drugą ideą było określenie wpływu uzyskanych czynników dla poszczególnych zespołów w końcowym rankingu Mistrzostw Świata w koszykówce w Chinach w 2019 roku.

1. PODSTAWY TEORETYCZNE PROBLEMU BADAWCZEGO

Literatura dotycząca problemu analizy gry w koszykówkę sukcesywnie się rozrasta, (szczególnie w liczbie możliwych ujęć problemu), tym samym, ustawicznie ukazując wciąż nowe perspektywy. Taki stan faktyczny tworzy doskonałą przestrzeń, aby w nowoczesny sposób rozwiązać sektor analizy strukturalnych elementów gry w koszykówkę i określić, jakie składowe determinują poziom mistrzostwa sportowego najlepszych na świecie reprezentacji narodowych.

Podstawowym wyzwaniem badaczy z całego świata jest odnalezienie sposobu rejestracji i analiz gry oraz zebrania nowych informacji diagnostycznych m.in. z zawodów

⁵ D. Zambová, L. Tománek, *An efficiency shooting program for youth basketball players*, „Sport Logia”, 2012, nr 8, s. 87-92.

⁶ R. R. Oudejans, S. Heubers, J. R. J. Ruitenbeek, T. W. Janssen, *Training visual control in wheelchair basketball shooting*, „Research Quarterly for Exercise and Sport”, 2012, vol 83, s. 464- 469.

⁷ K. Matulaitis, S. Stonkus, *Įvairaus amžiaus krepšinio komandų žaidimo rezultatyvumo ir metimų įvairovės rodiklių analizė*, „Baltic Journal of Sport and Health Sciences”, 2009, nr 1,

⁸ R. Khelifa, i wsp., *Kinematic adjustments in the basketball free throw performed with a reduced hoop diameter rim*, „International Journal of Sports Science & Coaching”, 2012, vol 7,

⁹ J. L. Arias, F. M. Argudo, J. I. Alonso, *Effect of basketball mass on shot performance among 9–11 year-old male players*, „International Journal of Sports Science & Coaching”, 2012, vol 7, s. 69-79.

rangi mistrzowskiej. Zaprezentowana w niniejszej dysertacji analiza danych, pochodzących z syntetycznej oceny przebiegu gry, poprzez zastosowanie zaawansowanych narzędzi statystycznych oraz stworzenie modelu teoretycznego, wysoko skorelowanego z rzeczywistym rankingiem Mistrzostw Świata w Chinach, stanowi wzór i kierunek myślowy dla poszerzania horyzontów nauki.

Wnikliwe analizy przebiegu gry oraz przegląd literatury stanowią konieczny warunek dla dokonania rzetelnej oceny efektywności gry w kontekście indywidualnym i zespołowym. Określenie czynników diagnostycznych, które w zestawieniu z rankingiem Mistrzostw Świata najbardziej charakteryzują poszczególne zespoły oraz tych, które mają największy wpływ na poziom mistrzostwa sportowego i finalny ranking, stanowi wyzwanie dla badaczy z całego świata.

2. CEL PRACY I PYTANIA BADAWCZE

Określone zostały dwa zasadnicze cele badań. Pierwszym z nich jest wyodrębnienie elementów strukturalnych gry determinujących mistrzostwo sportowe na poziomie czołowych światowych reprezentacji narodowych w koszykówce mężczyzn, podczas Mistrzostw Świata w 2019 roku. W tym kontekście niezwykle istotne wydaje się stworzenie metody, która rzetelnie i z dużą szczegółowością przeprowadzi analizę syntetyczną i dokona oceny wszystkich gier mistrzostw zespołów sklasyfikowanych w pierwszej ósemce rankingu. Celem analizy poszczególnych meczów będzie uzyskanie zmiennych diagnostycznych, które w istotny sposób poszerzają zakres obecnie stosowanych danych statystycznych w meczach koszykówki. Takie informacje, z dużym prawdopodobieństwem rzucają nowe światło do badań w zakresie analizy gry, a trenerom ułatwią właściwe planowanie i realizację treningu sportowego oraz rozłożenie akcentów treningowych i meczowych.

Drugim celem pracy jest wielowymiarowa analiza statystyczna determinująca wynik sportowy na najwyższym poziomie rywalizacji w koszykówce podczas Mistrzostw Świata w koszykówce mężczyzn w Chinach w 2019 roku. W tym kontekście niezwykle istotne wydaje się wyznaczenie tych elementów (zmiennych diagnostycznych), które charakteryzują grę zespołów oraz mają największy wpływ na zajmowane lokaty rankingowe. W oparciu o uzyskane rezultaty badań zostanie podjęta próba zbudowania syntetycznego modelu, który z dużym prawdopodobieństwem, na podstawie zmiennych diagnostycznych określi ranking zajmowanych miejsc zgodnie z rankingiem rzeczywistym mistrzostw.

Biorąc pod uwagę cele pracy sformułowano następujące pytania badawcze:

1. W jaki sposób przeprowadzić rejestracje walki sportowej i wstępną analizę statystyczną umożliwiającą wyodrębnienie zwiększonego zakresu elementów charakteryzujących grę?
2. Które zmienne diagnostyczne w największym stopniu determinują rzeczywisty ranking mistrzostw i powalają na przyporządkowanie zespołów do grup?
3. Czy na podstawie uzyskanych wyników istnieje możliwość zbudowania modelu teoretycznego, który jest silnie powiązany rankingiem rzeczywistym mistrzostw?
4. Czy na podstawie uzyskanych wyników można określić czynniki gry, które w największym stopniu prognozują prawdopodobieństwo zwycięstwa i porażki w koszykówce?

Powyższe pytania skłaniają do wysunięcia następujących hipotez badawczych:

H1. Istnieje możliwość stworzenia nowego systemu zapisu rejestracji gry oraz jej statystycznej analizy, który dostarczy nowe element diagnostyczne, które rozszerzają zakres powszechnie stosowanych obserwacji meczowych.

H2. Istnieją zmienne diagnostyczne związane ze skutecznością gry ofensywnej, które w największym stopniu determinują rzeczywisty ranking mistrzostw i powalają na przyporządkowanie zespołów do grup.

H3. Istnieje możliwość zbudowania modelu teoretycznego, który jest bardzo silnie powiązany rankingiem rzeczywistym mistrzostw.

H4. Na podstawie uzyskanych wyników można określić czynniki gry, które najlepiej prognozują prawdopodobieństwo zwycięstwa i porażki w koszykówce.

3. MATERIAŁ I METODY BADAŃ

Warunkiem wyznaczenia zespołów do badanej grupy było zajęcie czołowych lokat w Mistrzostwach Świata w koszykówce mężczyzn w Chinach w 2019 roku. Każdy z ośmiu zespołów w całym turnieju rozegrał osiem meczów. Następnie każdy z tych meczów został poddany analizie syntetycznej. Wyjątek stanowiła drużyna Australii, która została przeanalizowana na podstawie siedmiu rozegranych spotkań¹⁰. W celu uzyskania rzetelnych pomiarów, konieczne było zastosowanie takiego dostępu do meczów, który dawał wysoki standard wizualny i dźwiękowy. Wszystkie analizy meczów zostały wykonane za pomocą strony internetowej <https://www.eurosportplayer.com>, na podstawie zapisu audiowizualnego

¹⁰ Mecz Australia : Dominikana - prawa autorskie zostały zastrzeżone, mecz był zatem niedostępny i w ten sposób niemożliwy do zanalizowania.

sześćdziesięciu trzech meczów z Mistrzostw Świata. Wykorzystując funkcję stop-klatki, dane z przebiegu gry zostały zapisane w autorskim arkuszu obserwacji.

Aby sprostać kryterium rzetelności badań, każdy mecz został analizowany przez dwóch, niezależnych obserwatorów, co stanowiło sumę stu dwudziestu sześciu odrębnych analiz meczowych. Wyniki meczów zostały ze sobą porównane w ocenie spójności i rzetelności zapisów w siedmiu obszarach. Sumaryczny wynik rzetelności został wyrażony w ujęciu procentowym,

a różnica pomiędzy obserwatorami stanowiła liczbę poniżej 1%. W efekcie zebrano dane syntetyczne, umożliwiające podjęcie właściwych badań.

Powszechnie stosowanym narzędziem służącym do oceny struktury gry są statystyki, które są niezbędne w obiektywnej obserwacji zarówno poszczególnych zawodników, jak i całego zespołu. Tradycyjny arkusz statystyczny nie pozwala skutecznie ocenić pełnego spektrum przydatności zawodnika i zespołu w trakcie pojedynczego meczu, imprezy sportowej czy całego sezonu. Powszechnie stosowane arkusze obiektywnej obserwacji meczowej informują wyłącznie o statystykach ogólnych, bez uwzględnienia wahań jakościowego wyrażania działań zawodników i skuteczności wskaźników wydajności. Analizy dotyczące liczby asyst, zbiórek, przechwyty, bloków, rzutów, minut spędzonych na parkiecie przez zawodnika itp. tylko częściowo, w niepełny sposób umożliwiają wgląd do indywidualnej czy zespołowej oceny gry.

W niniejszej pracy, w pierwszym etapie badań została zastosowana autorska metoda obserwacji kontrolowanej, z oceną rzetelności obserwatorów oraz programem, dokonującym obliczeń arytmetycznych. Na poszczególnych etapach uzyskiwane dane były weryfikowane oceną ekspertów, a także poddawane analizie rzetelności, aby następnie móc wykonać obliczenia, przy zastosowaniu autorskiego programu obliczeniowego. Etap ostatni polegał na wykorzystaniu wielu różnych narzędzi analizy statystycznej, z uwzględnieniem macierzy kwalifikacji, porządkowania liniowego, analizy skupień czy krzywej ROC.

Niniejsza praca od początku swojej koncepcji miała charakter poszukiwania nowych zmiennych diagnostycznych, które tak jak wskazuje przegląd literatury, nie były rozpatrywane w większości badań w XXI wieku. Prognozowanym rezultatem niniejszych rozważań będzie określenie nowych, szczegółowych wskaźników zespołowych, umożliwiających wypracowanie liczbowego wzorca, który będzie określał rezultat niezbędny do osiągnięcia sukcesu sportowego w pojedynczych meczach oraz całych turniejach i sezonach.

3.2 Techniki i narzędzia badawcze

Algorytmy zastosowanego programu wyodrębniały obliczenia w trzech sektorach diagnostycznych. Pierwszy, składający się z trzydziestu trzech zmiennych diagnostycznych, dotyczący *analizy w zakresie podstawowym*. Druga grupa obliczeń została przyporządkowana jako *przedziały czasowe* i zawierała w sobie trzydzieści pięć wskaźników, związanych z czasem rozgrywanych akcji i tempem gry zespołów. Natomiast trzeci obszar obliczeń dotyczył *ciągów rządowych akcji*, który różnicował czterdzieści cztery wskaźniki gry, oparte o skuteczne i nieskuteczne sekwencje akcji zespołu. Łącznie, dzięki stworzonemu *software'owi* istnieje możliwość obliczania stu dziesięciu zmiennych diagnostycznych z pojedynczego meczu. Przy czym analizując liczbę działań ofensywnych obu zespołów, w każdej konfrontacji, naturalne jest otrzymanie wskazań dotyczących działań przeciwnika. Tym samym, przyjęcie faktu zawierającego wyniki działań rywala stanowi o skuteczności i wskaźnikach defensywnych każdego zespołu.

3.2.1 Autorski arkusz obserwacji przebiegu gry

Pierwszym etapem metodologicznym, koniecznym do uzyskania zmiennych, determinujących przebieg meczu koszykówki, było przygotowanie arkusza obserwacji, który pozwalał na zgromadzenie danych bezpośrednio z przebiegu gry. Do tego działania użyto programu Microsoft Excel, wraz z szeregiem zasad i poleceń niezbędnych do sporządzenia prawidłowego zapisu. Założono, że nazwa gromadzonych plików powinna pochodzić od nazw zespołów, podejmujących bezpośrednią rywalizację w mistrzostwach.

Tabela 1 prezentuje prawidłowy zapis pierwszej kwarty przykładowego meczu rozegranego podczas turnieju w Chinach w 2019 roku.

Tabela 1. Fragment arkusza obserwacji meczu – pierwsza kwarta.¹¹

kwarta	posiadanie	liczba akcji w posiadaniu	czas trwania akcji	kod akcji
1	1	1	6	2
1	2	1	11	0/2
1	3	3	15;1;12	F;0/3;2
1	4	1	16	2
1	5	1	6	2
1	6	1	13	2
1	7	1	10	2
1	8	1	1	2
1	9	1	5	BR
1	10	1	4	0/3
1	11	1	17	BR
1	12	1	13	0/2
1	13	1	15	0/3
1	14	1	11	0/2
1	15	1	7	0/3
1	16	1	3	2
1	17	1	14	2
1	18	1	17	BR
1	19	1	2	2
1	20	3	6;1;1	F;P;BR

3.2.2 Analiza rzetelności – porównanie obserwatorów

Drugim etapem w procesie metodologicznym było przygotowanie analizy ze wszystkich meczów mistrzostw. Odpowiedzialnością za ten fakt zostali obarczeni wykwalifikowani trenerzy z wieloletnim doświadczeniem pracy z drużynami koszykówki. Stosując metodę obserwacji zapisu video z funkcją stop-klatki, dwóch obserwatorów w niezależnych miejscach i ściśle określonym czasie dokonywali samodzielnej analizy powierzonych materiałów z zapisem meczów.

Zgodnie z definicją, rzetelność dotyczy spójności zapisu, czyli stopnia zgodności wyników, jakie można uzyskać za pomocą narzędzi i procedur, badając daną grupę

¹¹ Zapis kolejnych kwart i ewentualnych dogrywek w meczu powinien przebiegać analogicznie do powyższej tabeli.

czy populację w różnych okolicznościach. Rzetelność daje badaczom informację na temat stopnia przypadkowej niespójności oraz niesystematycznej fluktuacji indywidualnych wyników w pomiarze. Trzeba pamiętać, że teoretycznie trafny pomiar jest zawsze rzetelny, natomiast rzetelny pomiar nie zawsze musi być trafny. Inaczej mówiąc, narzędzie pomiaru może być rzetelne, ale nie musi w sposób trafny odzwierciedlać danego konstrukt¹².

W związku z powyższym, w niniejszej pracy dla spełnienia wymogów oceny rzetelności, uwzględniono porównanie zapisów dwóch obserwacji meczu, przygotowanych przez dwóch niezależnych obserwatorów dla wszystkich meczów mistrzostw. Skala porównania zawierała siedem płaszczyzn obserwacji (kolumn). Informacje zostały zapisane w postaci wartości różnicy procentowej pomiędzy zapisami oraz wartości liczbowej i wskazywały one ostateczne wyniki pomiaru rzetelności.

Szczegółowy skład czynnikowy analizy rzetelności prezentuje się następująco:

- 1. Błąd zapisu kwarty [%]** – różnica w numeracji kwarty między obserwatorami.
- 2. Błąd zapisu kolejności posiadania [%]** – różnica w numeracji posiadania między obserwatorami.
- 3. Błąd zapisu liczby akcji [%]** – różnica w zapisie liczby akcji w posiadaniu między obserwatorami.
- 4. Błąd zapisu czasu zakończenia akcji [%]** – różnica w zapisie czasów akcji między obserwatorami.
- 5. Błąd zapisu kodu akcji [%]** – różnica w zapisie kodów akcji między obserwatorami.
- 6. Błąd sumaryczny [%]** – suma wszystkich błędów w całym pliku bez podziału na kolumny.
- 7. Różna liczba zapisanych akcji** – w sytuacji gdy np. jeden obserwator dokonał zapisu trzech akcji w posiadaniu, a drugi obserwator tylko dwóch. Kolumna podaje liczbę wystąpień opisanej sytuacji w danym zapisie.

Tabela 2 przedstawia analizę rzetelności dwóch obserwatorów podczas porównania monitorowanych spotkań zespołu Argentyny. Pełen zapis analizy rzetelności dla wszystkich meczów Mistrzostw Świata w Chinach w 109 roku znajduje się w aneksie.

¹² A. Michońska-Stadnik, *Trafność i rzetelność w badaniach glottodydaktycznych*, 'Lingwistyka Stosowana' 2011, nr 4, s. 31-40.

Tabela 2. Arkusz analizy rzetelności obserwatorów – zespół Argentyny¹³.

l.p.	mecz	błąd - kwarty [%]	błąd - posiadania [%]	błąd - akcje [%]	błąd - czasy [%]	błąd - kody akcji [%]	błąd - sumaryczny [%]	różna liczba zapisanych akcji
1	argentyna_francja	-	-	-	2,20	2,20	1,03	-
2	argentyna_francja	-	-	-	1,96	0,98	0,72	-
3	argentyna_hiszpania	-	-	1,23	1,03	-	0,46	-
4	argentyna_hiszpania	1,22	-	-	-	0,94	0,44	-
5	argentyna_korea	-	-	-	-	-	-	-
6	argentyna_korea	-	-	-	-	-	-	-
7	argentyna_nigeria	-	-	-	-	-	-	-
8	argentyna_nigeria	-	-	-	-	-	-	-
9	argentyna_polska	-	-	-	2,97	-	0,66	-
10	argentyna_polska	-	-	-	-	3,77	0,86	-
11	argentyna_rosja	-	-	-	0,97	-	0,23	-
12	argentyna_rosja	-	-	-	-	-	-	-
13	argentyna_serbia	-	-	-	0,97	-	0,23	-
14	argentyna_serbia	-	-	2,50	2,68	0,89	1,29	-
15	argentyna_wenezuela	-	-	-	2,25	1,12	0,77	-
16	argentyna_wenezuela	-	-	-	9,57	1,06	2,51	-

Powyższe wyniki wskazują m.in. na brak różnic w różnicy w liczbie zapisanych akcji oraz na fakt, iż nie odnotowano błędów w zapisie kolejności posiadania. Błąd zapisu kwarty wystąpił jednokrotnie, błąd dotyczący liczby akcji w posiadaniu wystąpił dwukrotnie. Inne różnice określone jako błąd, pojawiały się częściej, jednak ich występowanie, wyrażone w wartości procentowej było minimalne.

Otrzymane wyniki ze wszystkich meczów Mistrzostw Świata w Chinach w zakresie rzetelności obserwacji prezentowały się następująco:

- Różnica zapisu kwarty:** w 126 meczach obserwatorzy różnili się w 9 zapisach numerów kwart. Najmniejsza różnica pomiędzy analizami meczów w tym wskaźniku wyniosła 1,22%, największa różnica wyniosła zaś 4.29%. Nie odnotowano różnic w 117 przypadkach. Średnia różnica kwarty wyniosła 0,14%.

¹³ Zapis kolejnych meczów przebiegał analogicznie.

2. **Różnica zapisu posiadania:** w 126 meczach obserwatorzy różnili się w 2 zapisach posiadania. Oba przedstawiają się na poziomie 1.32%. Nie odnotowano różnic w 124 przypadkach. Średnia różnica zapisu liczby posiadania wyniosła 0,02%.
3. **Różnica zapisu ilości akcji:** w 126 meczach obserwatorzy różnili się w 26 zapisach ilości akcji w posiadaniu. Najmniejsza różnica pomiędzy analizami meczów w tym wskaźniku wyniosła 1,06%, największa różnica wyniosła 4.11%. Nie odnotowano różnic w 100 przypadkach. Średnia różnica zapisu liczby akcji w posiadaniu wyniosła 0,34%.
4. **Różnica zapisu czasu zakończenia akcji:** w 126 meczach obserwatorzy różnili się w 60 zapisach czasów zakończenia akcji. Najmniejsza różnica pomiędzy analizami meczów w tym wskaźniku wyniosła 0,83%, największa różnica wyniosła 9.57%. Nie odnotowano różnic w 66 przypadkach. Średnia różnica zapisu liczby czasu zakończenia akcji wyniosła 1,11%.
5. **Różnica zapisu kodu akcji:** w 126 meczach obserwatorzy różnili się w 63 zapisach kodu akcji. Najmniejsza różnica pomiędzy analizami meczów w tym wskaźniku wyniosła 0,83%, największa różnica wyniosła 8,25%. Nie odnotowano różnic w 63 przypadkach. Średnia różnica liczby zapisu kodu akcji wyniosła 1,02%.
6. **Różnica sumaryczna:** w 126 meczach wynik sumaryczny obserwatorów różnił się w 87 meczach. Najmniejsza różnica pomiędzy analizami meczów w tym wskaźniku wyniosła 0,22%, największa różnica wyniosła 3,71%. Nie odnotowano różnic w 39 przypadkach. Średnia sumaryczna różnica między obserwatorami wyniosła 0,58%.
7. **Różna liczba zapisanych akcji:** w 126 meczach obserwatorzy różnili się w 1 zapisie różnej liczby zapisanych akcji. W jednym meczu różnica wynosiła 2%. Nie odnotowano różnic w 125 przypadkach. Średnia różnica liczby zapisu kodu akcji wyniosła 0,03%.

W oparciu o powyższe treści, po zestandaryzowaniu wyników, błąd sumaryczny dla wszystkich meczów Mistrzostw Świata wynosił zaledwie 0,58% pomiędzy dwoma obserwatorami. Zaistniały fakt jednoznacznie wskazał, że do badań właściwych można było zastosować analizy dowolnego obserwatora.

3.3 Modułowe narzędzie do rejestracji i analizy działań ofensywnych

Do przygotowania modułowego programu¹⁴ rejestracji i analiz gry w koszykówkę wykorzystano możliwości programu MATLAB, który jest doskonałym, interaktywnym środowiskiem do wykonywania obliczeń naukowych i inżynierskich, oraz do tworzenia symulacji komputerowych. W pierwszym kroku tworzenia narzędzia należało przypisać wartości numeryczne do danych uzyskanych z arkusza obserwacji. Następnie algorytmy programu zestawiały ze sobą szeregi liczb, tworząc odpowiednie podzbiory i zbiory danych.

3.3.1 Pierwszy moduł analiz – dane podstawowe

Pierwszym obszarem obliczeń w programie było wyodrębnienie podstawowych wskaźników gry w ofensywie. W tym przede wszystkim: liczba posiadania, liczba akcji, średnia liczba akcji na posiadanie, średnia liczba punktów na akcję, średnia liczba punktów na posiadanie i inne średnie wskazujące poziom skuteczności działań. W tym, szczegółowo:

1. **liczba rzutów wolnych** - liczba rzutów wolnych z gry, liczba rzutów wolnych po limicie fauli, liczba rzutów wolnych technicznych, liczba rzutów oddanych, celnych i niecelnych oraz ich skuteczność,
2. **liczba rzutów za 2 punkty** – liczba oddanych rzutów celnych i niecelnych oraz ich skuteczność,
3. **liczba rzutów za 3 punkty** – liczba oddanych rzutów celnych i niecelnych oraz ich skuteczność,
4. **liczba rzutów 2+1 i 3+1** – liczba oddanych rzutów celnych i niecelnych oraz ich skuteczność,
5. **liczba dobitek** – liczba oddanych rzutów celnych i niecelnych oraz ich skuteczność,
6. **liczba innych działań** – liczba fauli, przerwanych akcji i akcji bez rzutu.

W tabeli 15 zaprezentowano wyniki pierwszego modułu analizy. Zestawienie danych pochodzi z meczu reprezentacji Australii i Francji. Dane przedstawiają klasyczne ujęcie obserwacji meczowych, które często znajdują swoje zastosowanie w pracy trenerskiej oraz w badaniach naukowych.

¹⁴ Program stworzony przez Tomasza Wilczewskiego i Wojciecha Marszałka został opracowany jako wynik wieloletniej wymiany doświadczeń, która miała swój początek na studiach doktoranckich na Akademii Wychowani Fizycznego w Katowicach w 2018 roku. W tym okresie między młodymi badaczami pojawiły się koleżeńskie relacje, a w ich wyniku wielopoziomowa wymiana doświadczeń w zakresie sportu i programowania.

Tabela 3. Moduł danych podstawowych z meczu Australia – Francja.

Drużyna	punkty	akcje rzutowe	rzuty wolne (suma)	rzuty wolne z gry	rzuty wolne z limitu	rzuty wolne techniczne	oddane rzuty za 1 punkt	rzuty celne za 1 punkt	rzuty niecelne za 1 punkt	skuteczność za 1
Australia	59	60	6	4	1	1	9	5	4	0,55
Francja	67	72	6	4	2	0	13	8	5	0,61
oddane za 2 pkt	celne za 2 pkt	niecelne za 2 pkt	skuteczność za 2	oddane za 3 pkt	celne za 3 pkt	niecelne za 3 pkt	skuteczność za 3	akcje 2+1	akcje 3+1	dobitki
38	21	17	0,55	18	4	14	0,22	2	0	3
45	16	29	0,35	21	9	12	0,42	0	0	3
dobitki celne	dobitki niecelne	skuteczność dobitek	przerwania	faule	akcje bez rzutu	posiadania	akcje	akcje na posiadanie	pkt na akcje	pkt na posiadanie
1	2	0,33	2	5	19	72	86	1,19	0,68	0,81
1	2	0,33	3	7	14	73	96	1,31	0,69	0,91

Z powyższej tabeli wynika m.in., że zespół z Francji wykonał łącznie o 12 akcji rzutowych z gry więcej, niż reprezentacja Australii. Francuzi w całym meczu rozegrali również o 10 akcji więcej, a ich współczynnik zdobytych punktów na jedno posiadanie wyniósł 0.91, natomiast Australijczycy osiągnęli wskaźnik wynoszący 0,81 punktu na posiadanie. Dużą różnicę odnotowano również w skuteczności rzutów za 3-punkty, która wynosiła 20% na rzecz reprezentacji Francji.

3.3.2 Drugi moduł analiz – przedziały czasowe

Drugi moduł programu w swoim założeniu miał wskazywać, w jakim czasie zespoły biorące udział w Mistrzostwach Świata w Chinach rozgrywają swoje akcje. W tym celu określono siedem przedziałów czasowych¹⁵, w których szczegółowo określano liczbę działań zespołu. Uszczegóławiając, wyróżniono, w jakim przedziale czasowym zespoły zdobywają najwięcej punktów, w jakim zdobywają ich najmniej, oraz jakie jest średnie tempo gry dla pojedynczej kwarty, połowy i całego meczu.

¹⁵ Przedział czasowy trwania akcji – w rozumieniu autora przedział czasowy rozumie się jako wartość (wyrażana w sekundach) trwania akcji w posiadaniu od jego początku do zakończenia. Obliczenia różnic dokonuje się na podstawie obserwacji zegara 24 sekund. W badaniach zastosowano siedem przedziałów czasowych; poniżej 1 sek., 1-4 sek., 5-8 sek., 9-12 sek., 13-16 sek., 17-20 sek., 21-24 sek.

Całość wyników została wyrażona za pomocą następujących wskaźników:

1. liczba skutecznych i nieskutecznych akcji, przeprowadzonych w przedziale czasowym **poniżej 1 sekundy,**
2. liczba skutecznych i nieskutecznych akcji, przeprowadzonych w przedziale czasowym **1-4 sekundy,**
3. liczba skutecznych i nieskutecznych akcji, przeprowadzonych w przedziale czasowym **5-8 sekund,**
4. liczba skutecznych i nieskutecznych akcji, przeprowadzonych w przedziale czasowym **9-12 sekund,**
5. liczba skutecznych i nieskutecznych akcji przeprowadzonych w przedziale czasowym **13-16 sekund,**
6. liczba skutecznych i nieskutecznych akcji przeprowadzonych w przedziale czasowym **17-20 sekund,**
7. liczba skutecznych i nieskutecznych akcji przeprowadzonych w przedziale czasowym **21-24 sekund,**
8. średnia tempa rozgrywanych kwart, wyrażona w sekundach,
9. liczba punktów zdobytych przez zespół w poszczególnych kwartach,
10. średnia skuteczności zdobywanych punktów w poszczególnych kwartach.

W tabeli 4 zaprezentowano wyniki drugiego modułu analizy. Zestawienie danych pochodzi z meczu reprezentacji Australii i Francji. Dane przedstawiają liczbę rozegranych akcji skutecznych i nieskutecznych w ośmiu przedziałach czasowych.

Tabela 4. Moduł danych z przedziałów czasowych z meczu Australia – Francja.

drużyna	skuteczne 0 sek.	skuteczne 1-4 sek.	skuteczne 5-8 sek.	skuteczne 9-12 sek.	skuteczne 13- 16 sek.	skuteczne 17- 20 sek.	skuteczne 21- 24 sek.	nieskuteczne 0 sek.	nieskuteczne 1-4 sek.
Australia	0	4	3	7	6	2	6	0	9
Francja	0	8	5	5	7	1	4	0	7

	nieskuteczne 5-8 sek.	nieskuteczne 9-12 sek.	nieskuteczne 13-16 sek.	nieskuteczne 17-20 sek.	nieskuteczne 21-24 sek.	tempo kwarta 1	tempo kwarta 2	tempo kwarta 3	tempo kwarta 4
Australia	5	8	2	5	3	10,85	11,4	12,04	9,64
Francja	2	11	16	3	3	10,88	12	9,14	12,23
	-	skuteczność kwarta 1	skuteczność kwarta 2	skuteczność kwarta 3	skuteczność kwarta 4	pkt kwarta 1	pkt kwarta 2	pkt kwarta 3	pkt kwarta 4
Australia	-	0,5	0,43	0,5	0,37	16	14	16	13
Francja	-	0,29	0,23	0,45	0,58	11	10	21	25

Z powyższej tabeli wynika m.in., że zespół z Francji w przedziale czasowym 1-4 sek. wykonał o cztery skuteczne ataki więcej. Australijczycy zaś, w całym meczu rozegrali więcej skutecznych ataków w przedziałach czasowych 9-12 sek. i 21-24 sek. Średnie tempo rozgrywania poszczególnych kwart meczu w zespole z Europy zostało wyrażone na poziomie 10,88 sek., 12 sek., 9,14 sek. oraz 12,23 sek. W poszczególnych kwartach Francuzi osiągnęli skuteczność rzutową na poziomie 29%, 23%, 45%, 58%.

3.3.3 Trzeci moduł analiz – ciągi rządowe

Trzeci, innowacyjny obszar analizy został wyrażony jako wyniki zespołów w ofensywie w aspekcie skutecznych i nieskutecznych ciągów akcji¹⁶. W opisywanym ujęciu wyróżniono dwa rodzaje ciągów rządowych. Pierwszy z nich to ciągi skuteczne, składające się z akcji, w których zespół zdobywa co najmniej jeden punkt. Drugi rodzaj, to ciągi nieskuteczne, czyli takie, podczas których zespół w sekwencjach własnych działań nie był w stanie zdobyć nawet jednego punktu.

W tym obszarze wyróżnionych zostało osiem głównych sekcji:

1. liczba ciągów rządowych skutecznych w meczu,
2. liczba zdobytych punktów w meczu z ciągów rządowych,
3. średnia liczba zdobytych punktów w jednym rządowym ciągu akcji,
4. liczba akcji we wszystkich ciągach rządowych skutecznych,

¹⁶ Ciągi rządowe – w rozumieniu autora za ciąg uważa się co najmniej dwie akcje punktowe z rzędu. Przerwanie akcji lub faul bez rzutów wolnych nie przerywają ciągu akcji (zostają pominięte w badaniu). W przypadku zakończenia poszczególnych kwart, każdorazowo ciąg akcji jest zakończony.

5. średnia liczba akcji w pojedynczym ciągu rządowym skutecznym,
6. liczba ciągów rządowych nieskutecznych w meczu,
7. liczba akcji we wszystkich ciągach rządowych nieskutecznych,
8. średnia liczba nieskutecznych akcji w jednym rządowym ciągu.

W tabeli 5 zaprezentowano wyniki trzeciego modułu analizy. Zestawienie danych pochodzi z meczu reprezentacji Australii i Francji. Dane przedstawiają m.in. liczbę ciągów rządowych akcji, sumę zdobytych punktów w ciągach rządowych i sumę akcji skutecznych.

Tabela 5. Moduł danych z punktów rządowych z meczu Australia – Francja

Drużyna	ciągi rządowe skuteczne	suma zdobytych punktów	średnia punktów w jednym ciągu	suma akcji skutecznych	średnia liczba akcji skutecznych	liczba ciągów rządowych nieskutecznych	liczba akcji rządowych nieskutecznych	średnia nieskutecznych akcji w ciągu
Australia	5	25	5	12	2,4	13	41	3,15
Francja	6	34	5,6	16	2,66	15	50	3,33

Analiza porównawcza wykazała, iż zespół z Francji rozegrał o jedną sekwencję skutecznych akcji rządowych więcej od swoich przeciwników. W trakcie przeprowadzonych działań rządowych odnotowano dla zespołu z Australii dwanaście skutecznych akcji, w których zdobył on dwadzieścia pięć punktów. Natomiast zespół francuski w swoim rozkładzie akcji rządowych wykonał szesnaście skutecznych akcji, uzyskując wynik trzydziestu czterech punktów.

Dodatkowo, jak wskazuje tabela 6, ciągi rządowych akcji, zostały obliczone w sposób bardziej szczegółowy. Analiza wskazała podział rządowych akcji i liczbę ich występowania, zawierającą określoną liczbę zdobytych punktów przez każdy z zespołów.

Tabela 6. Moduł danych rzędowych dotyczący liczby zdobytych punktów

Drużyna	liczba rzędowych akcji z punktami: 2	liczba rzędowych akcji z punktami: 3	liczba rzędowych akcji z punktami: 4	liczba rzędowych akcji z punktami: 5	liczba rzędowych akcji z punktami: 6	liczba rzędowych akcji z punktami: 7	liczba rzędowych akcji z punktami: 8	liczba rzędowych akcji z punktami: 9	liczba rzędowych akcji z punktami: pow. 10
Australia	4	0	1	0	0	0	0	0	0
Francja	2	1	1	1	0	0	0	0	0

Specyfikacja ilościowa, dotycząca liczby akcji skutecznych w pojedynczym ciągu, wskazuje, iż zespół Australii w meczu rozgrywanym przeciwko Francji czterokrotnie rozegrał ciąg akcji skutecznych, trwający dwie akcje. Dodatkowo Australijczykom udało się rozegrać jeden ciąg akcji skutecznych, który składał się z czterech akcji rzędowych.

Kolejnymi danymi szczegółowymi analizowanymi przez program były ciągi akcji nieskutecznych, co zostało wyrażone w tabeli 7.

Tabela 7. Moduł danych rzędowych dotyczący liczby akcji bez punktów

Drużyna	liczba rzędowych akcji bez punktów: 2	liczba rzędowych akcji bez punktów: 3	liczba rzędowych akcji bez punktów: 4	liczba rzędowych akcji bez punktów: 5	liczba rzędowych akcji bez punktów: 6	liczba rzędowych akcji bez punktów: 7	liczba rzędowych akcji bez punktów: 8	liczba rzędowych akcji bez punktów: 9	liczba rzędowych akcji bez punktów pow. 10
Australia	5	4	2	1	1	0	0	0	0
Francja	6	3	3	1	2	0	0	0	0

Zebrane dane przedstawiają fakt, iż zespół z Australii pięciokrotnie nie był w stanie zakończyć akcji punktami serii dwóch akcji. Czterokrotnie ciąg akcji nieskutecznych trwał trzy akcje z rzędu. Dwukrotnie Francuzi powstrzymywali przeciwnika przez cztery kolejne akcje z rzędu a zespół Australii nie mógł pokonać obrony przeciwnika w pięciu i sześciu kolejnych seriach rzędowych akcji.

W programie bardzo przydatny okazał się wskaźnik, informujący o liczbie zdobytych punktów w ciągu, co zostało wyrażone w tabeli 8.

Tabela 8. Moduł danych rządowych dotyczący liczby zdobytych punktów

drużyna	ciąg rządowy poniżej 5 punktów	ciąg rządowy 6-7 pkt	ciąg rządowy 8-9 pkt	ciąg rządowy 10-11 pkt	ciąg rządowy 12-13 pkt	ciąg rządowy powyżej 14
Australia	0	4	0	1	0	0
Francja	3	1	1	1	0	0

Prezentowane dane w sposób klarowny wyraziły, że zespół Australii w tym meczu czterokrotnie wykonał sekwencję akcji, która pozwoliła zdobyć sześć lub siedem punktów. Ponadto jednokrotnie Australijczycy przełamali defensywę Francuzów dziesięcioma lub jedenastoma punktami w pojedynczej serii punktowej. Zespół z Europy charakteryzował się większą różnorodnością w liczbie zdobywanych punktów, zdobywając w serii akcji skutecznych trzykrotnie pięć lub mniej punktów i po jednym razie sześć lub siedem, osiem lub dziewięć oraz dziesięć lub jedenaście punktów.

Podsumowując zebrane informacje należy wspomnieć, że autorski program do wielowymiarowej analizy gry w koszykówkę został wielokrotnie sprawdzony w konwencjonalny sposób. Każdy obszar analiz na poszczególnych etapach był sprawdzany ręcznie przez co najmniej trzech ekspertów z dziedziny sportu i analizy. Pojawiające się niezgodności były wychwytywane i eliminowane w celu stworzenia doskonałej, bezbłędnej struktury programu. Dopiero po uzyskaniu całkowitej pewności i rzetelności analiz, zostały wykonane obliczenia właściwe, wyłuszczone w niniejszej pracy naukowej.

3.4 Narzędzia analizy statystycznej

W badaniach ocenom poddano cechy o charakterze ilościowym i jakościowym. Analiza takich danych ma swoją specyfikę, polegającą na zastosowaniu do porównań adekwatnych narzędzi statystycznych. Ze względu na bardzo dużą liczbę zmiennych w stosunku do liczby przypadków, w celu redukcji zbioru zmiennych, posłużono się metodami ekonometrycznymi. Aby wstępnie ograniczyć liczbę zmiennych diagnostycznych oraz dokonać identyfikacji zmiennych, powiązanych z rankingiem, w mistrzostwach użyto korelacji rang Spearmana (*Spearman rank correlation coefficient*)¹⁷, która jest jedną z nieparametrycznych miar

¹⁷ Ch. Spearman Ch., *The proof and measurement of association between two things*, „*American Journal of Psychology*”, 1904 n 15., s. 72–101.

monotonicznej zależności statystycznej między zmiennymi losowymi. Współczynnik ten jest wykorzystywany do opisu siły korelacji dwóch cech, wtedy gdy są one mierzalne, badana zbiorowość jest nieliczna oraz gdy mają charakter jakościowy i istnieje możliwość ich uporządkowania¹⁸. Miarę tę stosuję się również do badania zależności między cechami ilościowymi w przypadku niewielkiej liczby obserwacji¹⁹.

Kolejnym krokiem było wyznaczenie macierzy korelacji pomiędzy zmiennymi. Do dyskryminacji zmiennych zastosowano metodę odwróconej macierzy, która polega na wyznaczeniu macierzy odwrotnej do macierzy R - współczynników korelacji. Metodę odwróconej macierzy wykorzystano do wyboru zmiennych diagnostycznych do dalszych analiz.

W kolejnym etapie wykorzystano analizę dyskryminacyjną²⁰. Jej zastosowanie było konieczne do rozstrzygnięcia, które zmienne dyskryminują dwie lub więcej naturalnie wyłaniających się z grup przypadków. Narzędziem pozwalającym na to są funkcje (zmienne) dyskryminacyjne, określone na podstawie danych statystycznych dla obiektów, o których wiemy dokładnie, z jakich populacji pochodzą.

Kolejnym narzędziem analizy statystycznej wykorzystanym w pracy było porządkowanie liniowe Hellwiga²¹. Dla przeprowadzenia porządkowania liniowego, które cel stanowiło utworzenie syntetycznej zmiennej zależnej Y_{WR} – wskaźnik rozwoju według koncepcji Hellwiga²² lub w oparciu o metodę wzorca, w pierwszej kolejności zestandaryzowano wektor analizowanych zmiennych diagnostycznych²³.

¹⁸ M. Sobczyk, *Statystyka*, Wydawnictwo Naukowe PWN, Warszawa 2002.

¹⁹ W. Starzyńska, *Statystyka praktyczna*, Wydawnictwo Naukowe PWN, Warszawa 2005.

²⁰ Analiza dyskryminacyjna (ang. *discriminant analysis*) – zespół metod wielowymiarowej analizy danych. Zawiera metody, które czynią z tej techniki niezwykle efektywne narzędzie do zagadnień klasyfikacyjnych i eksploracji danych. Jej zadaniem jest rozstrzygnięcie, które zmienne niezależne (predyktory) w najlepszy sposób dzielą dany zbiór przypadków na występujące w naturalny sposób grupy, opisane jakościową zmienną zależną. Źródło: https://pl.wikipedia.org/wiki/Analiza_dyskryminacyjna dostęp: 1.12.2022

²¹ Metoda Hellwiga, zwana również metodą optymalnego wyboru predyktant lub metodą wskaźników pojemności informacji – formalna metoda doboru zmiennych objaśniających do modelu statystycznego (w szczególności modelu ekonometrycznego) stworzona w 1968 roku przez Zdzisława Hellwiga. Źródło: https://pl.wikipedia.org/wiki/Metoda_Hellwiga / dostęp: 2.12.2022

²² Z. Hellwig, *Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowanych kadr*, „Przegląd Statystyczny”, 1968, nr 4.

²³ Tenże, *Problem optymalnego wyboru predyktant*, „Przegląd Statystyczny”, 1969, nr 4.

Aby przeprowadzić porządkowanie liniowe, celem utworzenia syntetycznej zmiennej zależnej Y_{WR} – wskaźnik rozwoju według koncepcji Hellwiga²⁴ lub w oparciu o metodę wzorca, należy najpierw zestandaryzować wektor analizowanych zmiennych diagnostycznych²⁵. W pracy zastosowano również wskaźnik rozwoju zawodniczego Ryguły (WRZ)²⁶, bazujący na wskaźniku rozwoju Hellwiga.

Po redukcji liczby zmiennych zastosowano analizę dyskryminacyjną w oparciu o metodę krokową. W pracy wyegzemplifikowano również syntetyczną zmienną zależną w oparciu o metody porządkowania liniowego oparte na metodzie wzorca w postaci wskaźnika rozwoju zawodniczego Ryguły.

Kolejne analizy dotyczyły zagadnień taksonomicznych, a mianowicie analizy skupień. W badaniach sprawdzono, czy w oparciu o zmienne diagnostyczne jesteśmy w stanie wyodrębnić jednorodnie taksonomicznie podzbiory²⁷. Do analiz wykorzystano również analizę skupień bazującą na odległości Euklidesowej i metodzie Warda²⁸, która jest jedną z aglomeracyjnych metod grupowania, którą spośród pozostałych wyróżnia wykorzystanie podejścia analizy wariancji do oszacowania odległości między skupieniami. Zmierza ona do minimalizacji sumy kwadratów odchyłeń dowolnych dwóch skupień, które mogą zostać uformowane na każdym etapie. Uchodzi za metodę bardzo efektywną, mimo że prowadzi do tworzenia skupień o małej wielkości. Metoda Warda daje pełną kontrolę nad wynikową liczbą grup oraz przedstawia najbardziej naturalne skupiska elementów. Dopelnieniem analiz było wykorzystanie regresji logistycznej. Dla wszystkich analiz przyjęto poziom istotności równy 0,05.

W dalszych analizach zastosowano regresję logistyczną, która jest jedną z metod regresji, używanych w przypadku, gdy zmienna zależna (objaśniana) jest mierzona w skali dychotomicznej (przyjmuje tylko dwie wartości). Zwykle wartości zmiennej

²⁴ Z. Hellwig, *Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowanych kadr*, dz. cyt.,

²⁵ Tenże, *Problem optymalnego wyboru predyktant*, dz. cyt.

²⁶ I. Ryguła, *Elementy teorii, metodyki, diagnostyki i optymalizacji treningu sportowego*, AWF Katowice 2000.

²⁷ Taksonomia numeryczna - dyscyplina zajmująca się zasadami i procedurami porządkowania oraz klasyfikacji. Zadaniem metod taksonomicznych jest pogrupowanie zbioru elementów dowolnej natury Ω na bardziej jednorodnie statystycznie podzbiory, spełniające zarazem formalne warunki rozłączności, zupełności oraz niepustości. Źródło: https://mfiles.pl/pl/index.php/Metody_taksonomiczne /dostęp:4.12.2022

²⁸ Źródło: https://pl.wikipedia.org/wiki/Metoda_Warda dostęp:4.12.2022

objaśnianej wskazują na wystąpienie lub brak wystąpienia pewnego zdarzenia, które chcemy prognozować. Regresja logistyczna pozwala wówczas na obliczanie prawdopodobieństwa tego zdarzenia.

Kolejne analizy dotyczyły konstrukcji modeli wielowymiarowych na bazie regresji logistycznej dla zmiennej zależnych binarnych – przy czym klasą modelowaną było wystąpienie zwycięstwa czyli 1. Przy tworzeniu modelu wykorzystano V-krotną walidację krzyżową. Użyto testu Hosmera-Lemeshowa, który weryfikuje hipotezę o równości wartości obserwowanych i przewidywanych. Jeśli wartości obserwowane i przewidywane są wystarczająco bliskie, wówczas można założyć, że model jest dobrze dopasowany do danych. Brak istotności tego testu jest pożądanym. Jest to główna statystyka diagnostyczna modelu regresji logistycznej. Test Hosmera-Lemeshowa jest niezbędną statystyką podczas oceny dopasowania modelu do przewidywanych danych. Na podstawie tego testu rozpatruje się rewizję modelu, jego przekształcenie lub zastosowanie alternatywnej metody analizy²⁹.

W pracy, finalnie posłużono się narzędziem statystyki matematycznej jaką jest krzywa ROC (*receiver operating characteristic*), która jest graficzną reprezentacją efektywności modelu predykcyjnego, poprzez wykreślenie charakterystyki jakościowej klasyfikatorów binarnych, powstałych z modelu przy zastosowaniu wielu różnych punktów odcięcia. Finalnie na wykres nanosimy TPR (*True-Positive Rate* – oś pionowa) oraz FPR (*False-Positive Rate* – oś pozioma)³⁰.

Wszystkie analizy wykonano w oparciu o pakiet Statistica 13.1 oraz arkusz kalkulacyjny Excel.

4. WYNIKI BADAŃ

4.1 Redukcja liczby zmiennych – wybór zmiennych do badań

W przeprowadzonych badaniach wykorzystano 110 zmiennych, uzyskanych na podstawie analizy, jaką przeprowadzono z użyciem modułowego programu do rejestracji działań ofensywnych w koszykówce. Analizie zostały poddane zespoły zajmujące miejsca od 1. do 8. w rozegranych w Chinach Mistrzostwach Świata w koszykówce mężczyzn w 2019 roku.

²⁹ Źródło: <https://nauka.metodolog.pl/glossary/test-hosmera-lemeshowa/> dostęp: 5.12.2022

³⁰ Źródło: <https://mathspace.pl/matematyka/receiver-operating-characteristic-krzywa-roc-czyli-ocena-jakosci-klasyfikacji-czesc-7/> dostęp: 7.12.2022

W wyniku powyższych analiz, obliczono współczynniki korelacji rang Spearmana pomiędzy rankingiem rzeczywistym mistrzostw, a analizowanymi zmiennymi.

W poniższej tabeli przedstawiono tylko istotne, statystycznie powiązane z rankingiem zmienne. W związku z tym wstępna lista zmiennych diagnostycznych objęła dwadzieścia wskaźników. Zestawienie z rankingiem podano w tabeli 9.

Tabela 9. Korelacja porządku rang Spearmana

Para zmiennych	N Ważnych	R Spearman	t(N-2)	P
ranking & punkty	95	-0,20	-2,0	0,0472
ranking & rzuty wolne (suma)	95	-0,30	-3,1	0,0030
ranking & rzuty wolne z limitu	95	-0,42	-4,5	0,0000
ranking & oddane za 1 pkt	95	-0,29	-2,9	0,0050
ranking & celne za 1pkt	95	-0,28	-2,8	0,0058
ranking & akcje 3+1	95	-0,23	-2,3	0,0267
ranking & skuteczność dobitek	92	-0,22	-2,2	0,0332
ranking & skuteczne 1-4 sek.	95	-0,23	-2,3	0,0244
ranking & skuteczne 5-8 sek.	95	0,20	2,0	0,0500
ranking & skuteczne 17- 20 sek.	95	-0,22	-2,1	0,0356
ranking & skuteczne 21-24 sek.	95	-0,24	-2,4	0,0179
ranking & nieskuteczne 17-20 sek.	95	-0,25	-2,5	0,0157
ranking & nieskuteczne 21-24 sek.	95	0,21	2,1	0,0422
ranking & ra: 6	95	-0,21	-2,1	0,0428
ranking & pr: 12-13 pkt	95	-0,23	-2,3	0,0245
ranking & 2+0	95	0,20	2,0	0,0488
ranking & wl	95	-0,40	-4,2	0,0001
ranking & rządowe akcje bez pkt: 2	95	-0,21	-2,0	0,0452
ranking & rabp: 7	95	-0,25	-2,5	0,0148
ranking & rabp: 8	95	0,38	3,9	0,0002

Po określeniu i zgromadzeniu danych, dotyczących wstępnego zestawu zmiennych podejmuje się zazwyczaj w odniesieniu do nich odpowiednie działania weryfikacyjne według dwóch najistotniejszych kryteriów³¹. Istotna jest zmienność – zmienne powinny wykazywać odpowiednie zróżnicowanie, czyli skutecznie dyskryminować obiekty.

³¹A. Młodak, *Analiza taksonomiczna w statystyce regionalnej*, Wydawnictwo DIFIN, Warszawa 2006

Tabela 10. Współczynniki zmienności dla analizowanych zmiennych

Zmienna	Wsp.zmn. (%)
punkty	14,06
rzuty wolne (suma)	36,90
rzuty wolne z limitu	82,37
oddane za 1 pkt	37,45
celne za 1pkt	44,41
akcje 3+1	479,50
skuteczność dobitek	63,92
skuteczne 1-4 sek.	38,35
skuteczne 5-8 sek.	38,62
skuteczne 17-20 sek.	63,33
skuteczne 21-24 sek.	82,16
nieskuteczne 17-20 sek.	40,54
nieskuteczne 21-24 sek.	61,51
ra: 6	223,38
pr: 12-13 pkt	191,92
2+0	261,28
WL	100,24
Rzędowe akcje bez pkt: 2	40,37
rabp: 7	183,13
rabp: 8	266,40

Biorąc pod uwagę pierwsze z wyżej wymienionych kryteriów, żadnej z potencjalnych zmiennych diagnostycznych nie wyeliminowano z badania, gdyż obliczone dla nich współczynniki zmienności były bardzo wysokie (14% i więcej). Wystąpiła korelacja – dwie zmienne silnie ze sobą skorelowane są nośnikami podobnej informacji, a więc jedna z nich staje się zbędna.

4.2 Analiza zmiennych diagnostycznych – korelacje rang Spearmana

W tej sytuacji należy wziąć pod uwagę współczynniki korelacji wszystkich par zmiennych, a następnie zastosować odpowiednią metodę weryfikacji, w celu wyeliminowania zmiennych najbardziej podobnych do innych. Punktem wyjścia jest wyznaczenie macierzy korelacji zmiennych.

Tabela 11. Korelacje rang Spearmana dla analizowanych zmiennych diagnostycznych

	punkty	rzuty wolne (suma)	rzuty wolne z limitu	oddane za 1pkt	akcje 3+1	skuteczność dobietek	skuteczne 1-4 sek.	skuteczne 5-8 sek.	skuteczne 17-20 sek.	skuteczne 21-24 sek.	niekuteczne 17-20 sek.	niekuteczne 21-24 sek.	ra: 6	pr: 12-13 pkt	2+0	WL	rzędowe akcje bez pkt. 2	rabp: 7	rabp: 8
punkty	1,00	0,23	0,15	0,20	0,10	0,19	0,36	0,23	0,26	-0,06	-0,08	-0,31	0,21	0,42	0,08	0,21	0,18	-0,29	-0,03
rzuty wolne (suma)	0,23	1,00	0,65	0,97	0,18	0,28	0,29	-0,10	0,22	0,02	-0,02	-0,01	0,04	0,02	-0,02	0,69	0,02	0,14	-0,18
rzuty wolne z limitu	0,15	0,65	1,00	0,68	0,03	0,33	0,25	-0,27	0,20	-0,06	0,10	0,16	0,01	0,05	-0,13	0,87	0,12	0,18	-0,15
oddane za 1 pkt.	0,20	0,97	0,68	1,00	0,16	0,28	0,28	-0,16	0,17	0,00	-0,01	-0,03	0,02	-0,01	-0,14	0,70	-0,01	0,13	-0,16
celne za 1 pkt.	0,26	0,93	0,66	0,95	0,14	0,32	0,34	-0,17	0,18	-0,06	0,05	-0,05	0,01	0,03	-0,16	0,70	0,06	0,09	-0,13
akcje 3+1	0,10	0,18	0,03	0,16	1,00	0,08	-0,06	-0,01	0,24	0,20	0,18	-0,10	0,19	-0,12	-0,08	0,10	0,01	-0,12	-0,08
skuteczność dobietek	0,19	0,28	0,33	0,28	0,08	1,00	0,42	-0,02	0,00	-0,32	0,16	-0,09	0,11	0,21	0,08	0,19	0,03	0,16	-0,04
skuteczne 1-4 sek.	0,36	0,29	0,25	0,28	-0,06	0,42	1,00	0,08	-0,13	-0,23	0,11	-0,17	-0,08	-0,07	-0,10	0,22	0,16	0,09	0,06
skuteczne 5-8 sek.	0,23	-0,10	-0,27	-0,16	-0,01	-0,02	0,08	1,00	-0,28	-0,29	-0,31	-0,08	0,27	0,22	0,13	-0,29	0,09	-0,05	0,19
skuteczne 17-20 sek.	0,26	0,22	0,20	0,17	0,24	0,00	-0,13	-0,28	1,00	0,15	0,04	0,06	0,00	0,08	-0,02	0,28	0,08	-0,14	-0,22
skuteczne 21-24 sek.	-0,06	0,02	-0,06	0,00	0,20	-0,32	-0,23	-0,29	0,15	1,00	0,05	0,13	-0,01	0,00	-0,11	0,03	-0,12	-0,04	-0,20
niekuteczne 17-20 sek.	-0,08	-0,02	0,10	-0,01	0,18	0,16	0,11	-0,31	0,04	0,05	1,00	0,00	0,06	-0,09	0,01	-0,06	0,05	0,26	-0,08
niekuteczne 21-24 sek.	-0,31	-0,01	0,16	-0,03	-0,10	-0,09	-0,17	-0,08	0,06	0,13	0,00	1,00	-0,21	-0,19	0,03	0,07	-0,09	0,06	0,09
ra: 6	0,21	0,04	0,01	0,02	0,19	0,11	-0,08	0,27	0,00	-0,01	0,06	-0,21	1,00	0,41	-0,02	-0,01	-0,10	0,02	-0,18
pr: 12-13 pkt	0,42	0,02	0,05	-0,01	-0,12	0,21	-0,07	0,22	0,08	0,00	-0,09	-0,19	0,41	1,00	0,19	0,00	0,20	-0,05	0,00
2+0	0,08	-0,02	-0,13	-0,14	-0,08	0,08	-0,10	0,13	-0,02	-0,11	0,01	0,03	-0,02	0,19	1,00	-0,22	0,05	0,00	-0,07
WL	0,21	0,69	0,87	0,70	0,10	0,19	0,22	-0,29	0,28	0,03	-0,06	0,07	-0,01	0,00	-0,22	1,00	0,15	0,03	-0,11
rzędowe akcje bez pkt: 2	0,18	0,02	0,12	-0,01	0,06	0,03	0,16	0,09	0,08	-0,12	0,05	-0,09	-0,10	0,20	0,05	0,15	1,00	-0,14	-0,08
rabp: 7	-0,29	0,14	0,18	0,13	-0,12	0,16	0,09	-0,05	-0,14	-0,04	0,06	0,06	0,02	-0,05	0,00	0,03	-0,14	1,00	-0,08
rabp: 8	-0,03	-0,18	-0,15	-0,16	-0,13	-0,04	0,06	0,19	-0,22	-0,20	-0,08	0,09	-0,18	0,00	-0,07	-0,11	-0,08	-0,08	1,00

Tabela 12. Macierz odwrotna do macierzy współczynników korelacji pomiędzy zmiennymi diagnostycznymi

	Punkty	rzuty wolne (suma)	rzuty wolne z limitu	oddane za 1pkt.	celne za 1 pkt.	akcje 3+1	skuteczność dobielek	skuteczne 1-4 sek.	skuteczne 5-8 sek.	skuteczne 17-20 sek.	skuteczne 21-24 sek.	nieskuteczne 17-20 sek.	nieskuteczne 21-24 sek.	ra: 6	pr: 12-13 pkt	2+0	WL	rabp: 2	rabp: 7	rabp: 8
Punkty	2.46	1.35	-0.01	-0.41	-0.99	-0.08	0.15	-0.99	-0.76	-0.77	-0.24	-0.09	-0.47	-0.01	-0.42	-0.45	-0.42	0.21	0.58	0.02
rzuty wolne (suma)	1.35	41.08	3.56	-35.07	-4.63	-0.54	-0.19	-1.93	-3.45	-3.10	-2.20	0.37	-0.71	0.09	-0.42	1.35	1.35	0.21	0.58	0.02
rzuty wolne z limitu	-0.01	3.56	6.60	-5.38	1.48	0.27	-0.66	-1.93	-3.45	-3.10	-2.20	0.37	-0.71	0.09	-0.42	1.35	1.35	0.21	0.58	0.02
oddane za 1 pkt	-0.41	-35.07	-5.38	43.89	-8.66	0.07	0.19	1.79	-0.12	-0.01	0.38	-0.85	-0.94	-0.48	4.03	-0.41	-0.41	0.21	0.58	0.02
celne za 1 pkt	-0.99	-4.63	1.48	-8.66	14.29	0.20	-0.16	-0.20	1.03	0.53	1.39	-1.17	-0.47	0.68	1.25	-0.99	-0.99	0.21	0.58	0.02
akcje 3+1	-0.08	-0.54	0.27	-0.66	0.07	0.20	-0.41	0.34	-0.28	-0.27	-0.38	0.35	-0.47	0.31	0.10	-0.08	-0.08	0.21	0.58	0.02
skuteczność dobielek	0.15	-0.19	-0.66	0.19	-0.16	1.78	-0.70	0.27	0.02	0.02	0.54	-0.01	-0.04	0.03	-0.10	0.15	0.15	0.21	0.58	0.02
skuteczne 1-4 sek.	-0.99	-1.93	-0.18	1.79	-0.20	0.34	-0.70	2.08	0.02	0.52	0.04	-0.16	0.17	0.15	0.10	-0.99	-0.99	0.21	0.58	0.02
skuteczne 5-8 sek.	-0.76	-3.45	-0.12	2.16	1.03	-0.28	0.27	2.38	0.02	0.75	0.74	0.02	0.17	0.15	0.10	-0.76	-0.76	0.21	0.58	0.02
skuteczne 17-20 sek.	-0.77	-3.10	-0.01	2.48	0.53	-0.27	0.02	1.77	0.02	0.75	0.74	0.02	0.17	0.15	0.10	-0.77	-0.77	0.21	0.58	0.02
skuteczne 21-24 sek.	-0.24	-2.20	0.38	0.78	1.39	-0.38	0.54	0.04	0.75	0.74	1.76	-0.12	0.33	0.31	0.31	-0.24	-0.24	0.28	0.09	1.32
nieskuteczne 17-20 sek.	-0.09	0.37	-0.85	0.80	-1.17	-0.35	-0.01	-0.16	0.64	-0.01	1.67	0.00	0.00	0.00	-0.15	-0.09	-0.09	0.28	0.09	1.32
nieskuteczne 21-24 sek.	0.47	-0.71	-0.94	1.32	-0.47	0.19	-0.04	0.17	-0.33	-0.33	1.46	0.00	0.00	0.00	0.32	0.47	0.47	0.28	0.09	1.32
ra: 6	-0.01	-0.84	-0.05	0.68	0.31	-0.31	0.03	0.15	-0.46	0.16	1.71	-0.34	0.19	0.19	0.32	-0.01	-0.01	0.28	0.09	1.32
pr: 12-13 pkt	-0.77	0.09	-0.48	0.68	-0.82	0.56	-0.60	0.71	-0.11	-0.08	2.17	0.27	0.22	0.22	0.24	-0.77	-0.77	0.28	0.09	1.32
2+0	-0.42	-5.42	-0.45	4.03	1.25	0.10	-0.10	0.44	0.42	0.46	0.48	-0.15	-0.05	0.32	0.24	-0.42	-0.42	0.28	0.09	1.32
WL	-0.45	-4.32	-5.34	4.29	-1.39	-0.19	0.52	0.20	1.00	-0.03	-0.15	1.12	0.37	-0.22	0.51	0.82	0.82	0.28	0.09	1.32
rabp: 2	0.21	-0.35	-0.11	1.05	-0.40	-0.15	0.22	-0.27	-0.32	-0.04	0.19	-0.25	0.12	0.46	-0.51	-0.42	-0.42	0.28	0.09	1.32
rabp: 7	0.58	-1.58	-0.56	1.02	0.38	0.21	-0.05	-0.15	0.08	0.08	0.04	-0.41	0.15	0.07	-0.17	0.37	0.37	0.28	0.09	1.32
rabp: 8	0.02	0.22	0.36	0.02	-0.03	-0.13	0.10	-0.12	-0.21	0.23	0.31	-0.11	-0.19	0.41	-0.33	0.15	-0.42	0.28	0.09	1.32

Z tabeli 12 wynika, że należy wyeliminować trzy zmienne, dla których elementy diagonalne były wyższe od 10 (zaznaczone w tabeli kolorem czerwonym). Po ich eliminacji otrzymano 17 zmiennych diagnostycznych.

4.3 Klasyfikacja zespołów do grup – analiza dyskryminacyjna

Badane zespoły podzielono na cztery grupy, o zbliżonej liczebności. Na podstawie końcowego rankingu dokonano podziału koszykarskich drużyn biorących udział w Mistrzostwach Świata na cztery grupy :

1. Finałiści (miejsca 1- 2)
2. Grający o 3- 4 miejsce
3. Grający o 5- 6 miejsce
4. Grający o 7- 8 miejsce

Przy pomocy analizy dyskryminacyjnej dokonano próby rozpoznania tych czterech grup, w oparciu o wyegzemplifikowany zestaw zmiennych diagnostycznych, wykorzystując przy tym wszystkie zmienne uzyskane w oparciu o korelacje i macierz odwróconą. Po zastosowaniu krokowej procedury doboru zmiennych otrzymujemy ostateczny model analizy dyskryminacyjnej, który został przedstawiony w tabeli 13.

Tabela 13. Ostateczny model analizy dyskryminacyjnej przy wykorzystaniu 6 zmiennych dla analizowanych drużyn w oparciu o rozegrane przez nie mecze

N=92	Podsumowanie analizy funkcji dyskryminacyjnej. 6;Grupująca: Ranking (4 grup)					
	Lambda Wilksa: 0,42 przyb. F (18,235)=4,73; p< 0,00001					
	Lambda Wilksa	Cząstk. Wilksa	F usun. (3,83)	P	Toler.	1-Toler. (R- kwad.)
rzuty wolne z limitu	0,485	0,861	4,470	0,006	0,918	0,082
ra: 6	0,501	0,832	5,570	0,002	0,921	0,079
skuteczne 21- 24 sek.	0,543	0,769	8,329	0,000	0,712	0,288
rzędowe akcje bez pkt: 2	0,513	0,814	6,341	0,001	0,787	0,213
rabp: 7	0,479	0,872	4,058	0,010	0,880	0,120
skuteczność dobitek	0,471	0,885	3,587	0,017	0,796	0,204

Tabela 14. Macierz klasyfikacji modelu analizy dyskryminacyjnej dla analizowanych drużyn

Grupa	Procent Poprawne	Finał p=,26087	3 do 4 p=,23913	5 do 6 p=,26087	7 do 8 p=,23913
Finał (miejsca 1-2)	70,83	17,00	6,00	1,00	0,00
Miejsca 3 do 4	59,09	4,00	13,00	4,00	1,00
Miejsca 5 do 6	58,33	1,00	4,00	14,00	5,00
Miejsca 7 do 8	40,91	2,00	1,00	10,00	9,00
Razem	57,61	24,00	24,00	29,00	15,00

Analizując tabelę 14, wiersze odpowiadają klasyfikacji rzeczywistej, natomiast kolumny – klasyfikacji uzyskanej z wykorzystaniem 6 zmiennych. Klasyfikator jest idealny wtedy, gdy tylko na głównej przekątnej są liczby różne od zera. W naszym przypadku model jest dobry, gdy większość drużyn w oparciu o zmienne diagnostyczne uzyskane podczas rozegranych meczów została zaklasyfikowana zgodnie ze swoim statusem rzeczywistym. Można zauważyć, że utworzony model lepiej klasyfikuje finalistów w oparciu o zmienne diagnostyczne dla rozegranych przez nie meczów, przy czym poprawność klasyfikacji wynosi 70,83%. Wraz ze spadkiem w rankingu poprawność klasyfikacji się pogarsza i w przypadku drużyn z miejsc 7-8 wynosi 40,91%. Ogólna poprawność klasyfikacji na podstawie zmiennych diagnostycznych jest dobra, gdyż wynosi 57,61%.

Moduł analizy dyskryminacyjnej pozwala na wyliczenie funkcji klasyfikacyjnych, które mogą być wykorzystane do przyporządkowania drużyn do utworzonych grup, w oparciu o statystyki meczowe dla zmiennych diagnostycznych.

Tabela 15. Funkcje klasyfikacyjne – dla zmiennej grupującej ranking

Zmienna	Finał p=,26087	3 do 4 p=,23913	5 do 6 p=,26087	7 do 8 p=,23913
rzuty wolne z limitu	0,85	0,46	0,43	0,19
ra: 6	7,34	4,50	3,26	4,21
skuteczne 21-24 sek.	3,93	3,70	2,68	2,61
rzędowe akcje bez pkt: 2	2,53	2,30	1,79	1,79
rabp: 7	6,66	4,64	3,44	3,87
skuteczność dobitek	11,49	10,82	7,74	8,41
Stała	-21,10	-16,26	-9,88	-9,82

W tabeli 15 przedstawione drużyny zostały sklasyfikowane według największego prawdopodobieństwa w danym wierszu. Prawdopodobieństwa w główce tablicy oznaczają, że w zbiorze uczącym mieliśmy 26,09% obiektów grupy 1, 23,91% obiektów z grupy 2, 26,09% obiektów z grupy 3 oraz 23,91% obiektów z grupy 4.

Najwyższy wynik z tych wzorów wskazuje nam, do której z grup powinna przynależeć drużyna. Wprowadzono wartości 6 zmiennych diagnostycznych. Jako optymalny podzbiór zmiennych (ze zbioru przyjętego na początku), pozwalający dyskryminować (dzielić) drużyny na 4 grupy, w sposób podobny jak to stało się w zbiorze uczącym. Taką samą analizę dyskryminacyjną przeprowadzono również dla podziału na dwie grupy: strefa medalowa oraz miejsca 5-8. Przy pomocy analizy dyskryminacyjnej dokonano próby rozpoznania tych czterech grup w oparciu o wyegzemplifikowany zestaw zmiennych diagnostycznych, wykorzystując wszystkie zmienne uzyskane w oparciu o korelacje i macierz odwróconą. Po zastosowaniu krokowej procedury doboru zmiennych, otrzymujemy ostateczny model analizy dyskryminacyjnej, który został przedstawiony w tabeli 16.

Tabela 16. Ostateczny model analizy dyskryminacyjnej przy wykorzystaniu 7 zmiennych dla analizowanych drużyn w oparciu o rozegrane przez nie mecze

N=92	Podsumowanie analizy funkcji dyskryminacyjnej. 7;Grupująca: Ranking (2 grup)					
	Lambda Wilksa: ,52367 przyb. F (7,84)=10,915 p< 0,00001					
	Lambda Wilksa	Cząstk. Wilksa	F usun. (1,84)	p	Toler.	1-Toler. (R- kwad.)
rabp: 8	0,557	0,940	5,35	0,023	0,969	0,031
rzuty wolne z limitu	0,558	0,939	5,47	0,022	0,860	0,140
rzędowe akcje bez pkt: 2	0,597	0,878	11,72	0,001	0,830	0,170
skuteczne 21-24 sek.	0,642	0,815	19,05	0,000	0,678	0,322
skuteczność dobitek	0,584	0,897	9,67	0,003	0,788	0,212
nieskuteczne 21-24 sek.	0,562	0,931	6,19	0,015	0,916	0,084
rabp: 7	0,553	0,947	4,74	0,032	0,891	0,109

Tabela 17. Macierz klasyfikacji modelu analizy dyskryminacyjnej dla analizowanych drużyn

Grupa	Procent Poprawne	1 do 4 p=,50000	5 do 8 p=,50000
miejsca 1 do 4	82,61	38	8
miejsca 5 do 8	84,78	7	39
Razem	83,70	45	47

Analizując tabelę 17, wiersze odpowiadają klasyfikacji rzeczywistej, natomiast kolumny – klasyfikacji uzyskanej z wykorzystaniem 7 zmiennych. Klasyfikator jest idealny wtedy, gdy tylko na głównej przekątnej są liczby różne od zera. W przedstawionym przypadku model jest bardzo dobry, gdyż większość drużyn w oparciu o zmienne diagnostyczne uzyskane podczas rozegranych meczów została zaklasyfikowana zgodnie ze swoim statusem rzeczywistym do grup nazwanych jako „strefa medalowa” lub „miejsca 5–8”. Dla grupy ze strefy medalowej 1–4 poprawność klasyfikacji 82,61% natomiast dla grupy dotyczącej miejsc 5–8 wynosi 84,78%. Ogólna poprawność klasyfikacji na podstawie zmiennych diagnostycznych jest bardzo dobra, gdyż wynosi 83,70%.

Moduł analizy dyskryminacyjnej pozwala również na wyliczenie funkcji klasyfikacyjnych. Mogą one być wykorzystane do przyporządkowania drużyn do utworzonych grup, w oparciu o statystyki meczowe dla zmiennych diagnostycznych.

Tabela 18. Funkcje klasyfikacyjne – dla zmiennej grupującej ranking

Zmienna	Strefa medalowa p=,50000	5 do 8 p=,50000
rabp: 8	1,27	2,991
rzuty wolne z limitu	0,40	0,076
rzędowe akcje bez pkt: 2	2,17	1,658
skuteczne 21-24 sek.	3,50	2,395
skuteczność dobitek	10,92	7,816
nieskuteczne 21-24 sek.	0,35	0,765
rabp: 7	4,37	2,781
stała	-15,74	-9,782

Kolejnym etapem analiz było przeprowadzenie badania w oparciu o wyegzemplifikowane zmienne diagnostyczne porządkowania liniowego, celem uzyskania syntetycznego wskaźnika rozwoju 1.

4.4 Ranking teoretyczny i rzeczywisty – porządkowanie liniowe Hellwiga

Przed przeprowadzeniem porządkowania liniowego wykorzystano 17 zmiennych diagnostycznych, wyegzemplifikowanych przy pomocy korelacji i odwróconej macierzy. Na podstawie wysegregowanych zmiennych zbudowano teoretyczne zmienne endogeniczne Y_{WRZ}

dla uśrednionych wyników zmiennych diagnostycznych z wszystkich meczów rozegranych przez analizowane drużyny.

Tabela 19. Ranking – rzeczywisty z Mistrzostw Świata oraz Syntetyczny w oparciu o Y_{WR}

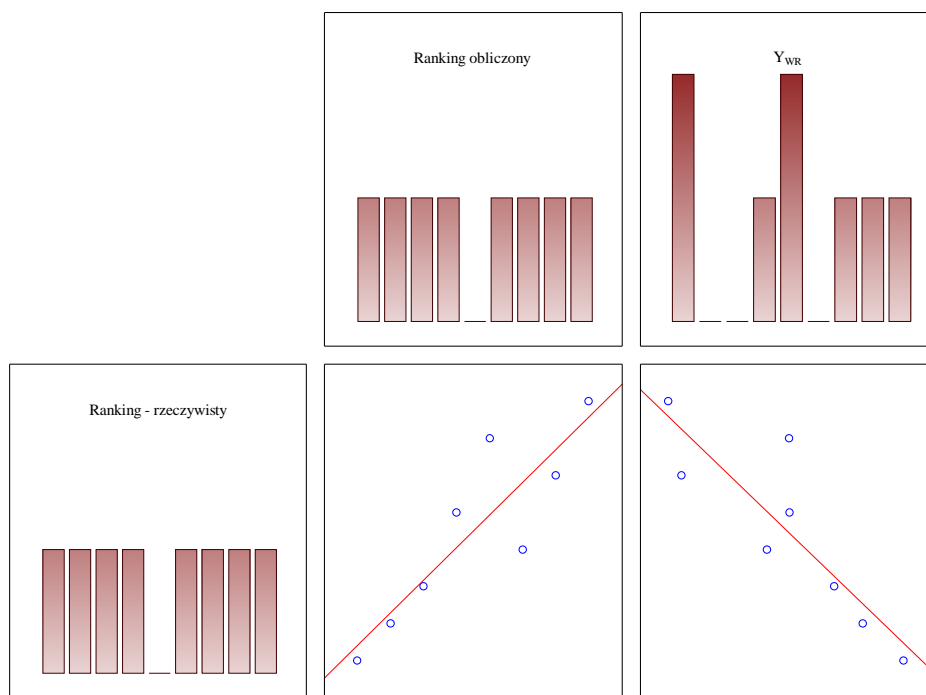
Drużyna	Ranking – rzeczywisty	Ranking obliczony – teoretyczny	Y_{WR}
Hiszpania	1	1	0,615
Argentyna	2	2	0,535
Francja	3	3	0,479
Australia	4	6	0,348
Serbia	5	4	0,392
Czechy	6	7	0,181
USA	7	5	0,391
Polska	8	8	0,155

Analiza wyników tabeli 19 wskazuje dosyć dużą zgodność pomiędzy rankingiem rzeczywistym i syntetycznym. Aby zweryfikować siłę powiązania, w dalszych analizach zastosowano współczynnik korelacji rang Spearmana.

Tabela 20. Korelacja porządku rang Spearmana

Para zmiennych	N Ważnych	R Spearman	t(N-2)	P
Ranking - rzeczywisty & Ranking obliczony na podstawie Y_{WR}	8	0,88	4,56	0,0039
Ranking - rzeczywisty & Y_{WR}	8	-0,88	-4,56	0,0039

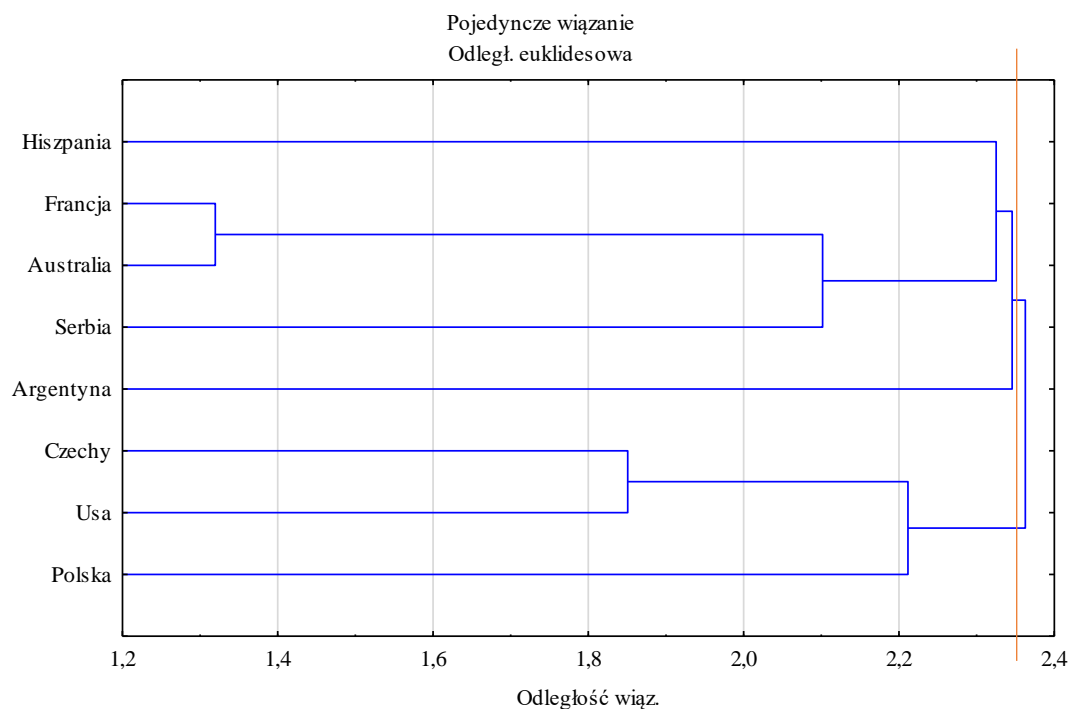
Analiza wyników pozwoliła na stwierdzenie, że ranking rzeczywisty i teoretyczny są bardzo silnie istotnie powiązane ze sobą $R=0,88$; $p0,0039$. Oczywiście analogicznie tak samo tylko ujemnie jest powiązany z rankingiem rzeczywistym syntetyczny wskaźnik rozwoju. Ujemna korelacja wynika z faktu, że zmienna rankingu jest destymulantą, czyli im niższy wynik, tym lepszy. Wyniki te potwierdza również poniższa rycina.



Rysunek 1. Korelacyjny macierzowy wykres rozrzutu

4.5 Jednorodne podzbiory mistrzostw – analiza taksonomiczna

Kolejne analizy dotyczyły zagadnień taksonomicznych, a mianowicie analizy skupień. W badaniach sprawdzono, czy w oparciu o zmienne diagnostyczne jesteśmy w stanie wyodrębnić jednorodne taksonomicznie podzbiory, zawierające analizowane drużyny.



Rysunek 2. Dendrogram – analiza skupień

Do analizy skupień wykorzystano odległość euklidesową oraz metodę Warda, w oparciu o które uzyskano dendrogram przedstawiony na rycinie 2. Wskazuje on, że w oparciu o dobrane zmienne diagnostyczne uzyskano dwa jednorodne taksonomicznie podzbiory przy punkcie odcięcia równym około 2,35. Stwierdzono, że podobne do siebie ze względu na wyniki zmiennych diagnostycznych były Hiszpania, Francja, Australia, Serbia oraz Argentyna. Drugi podzbiór reprezentowały Czechy, Polska oraz USA, co jest zgodne również z rzeczywistym rankingiem, w którym te drużyny zajęły miejsca 6-8.

4.6 Prognozy zwycięstwa i porażki – regresja logistyczna

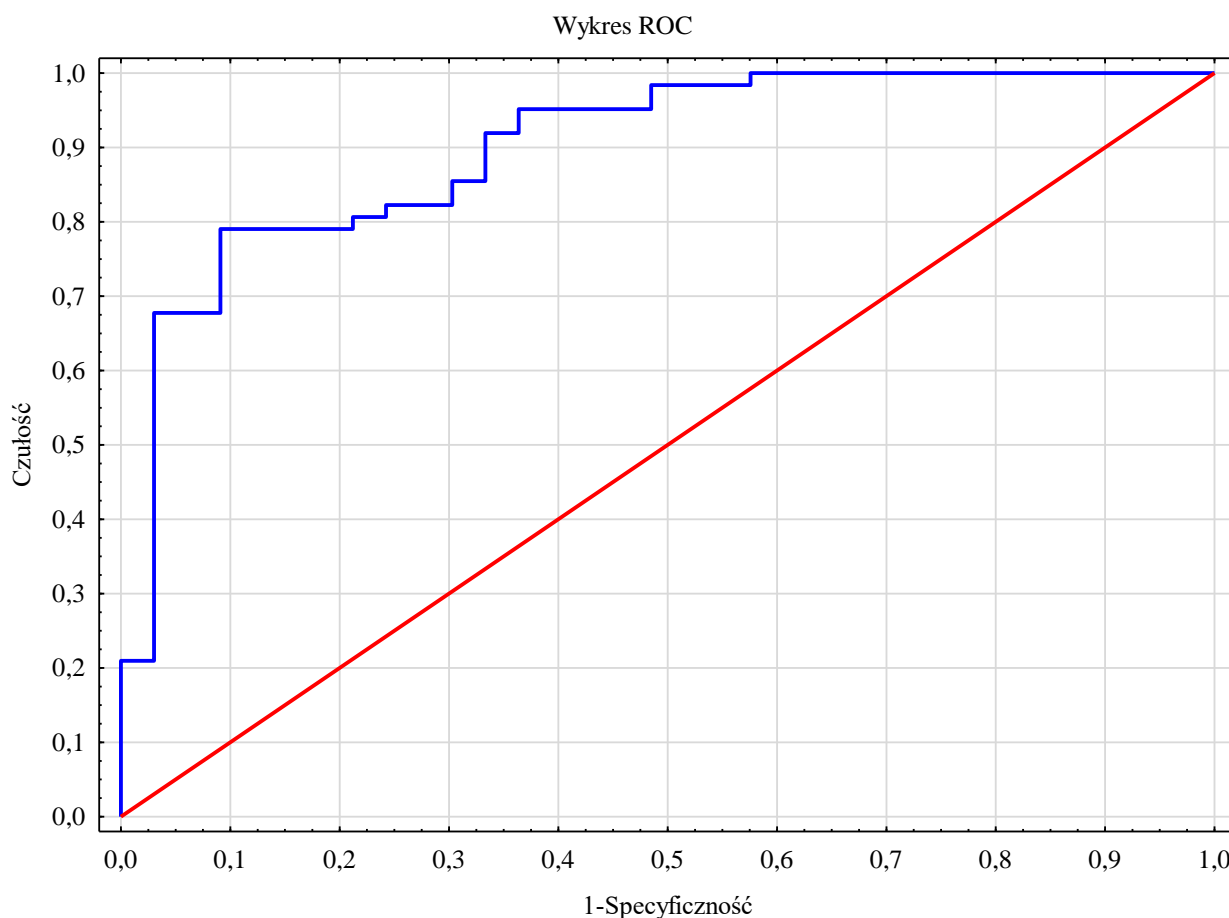
Kolejne analizy dotyczyły konstrukcji modeli wielowymiarowych na bazie regresji logistycznej dla zmiennej zależnych binarnych – przy czym klasą modelowaną było wystąpienie zwycięstwa czyli 1. Przy tworzeniu modelu wykorzystano V-krotną walidację krzyżową. Wyniki zamieszczono w poniższej tabeli.

Tabela 21. Z/P - Oceny parametrów. Rozkład: ranking DWUMIANOWY, F. wiążąca: LOGIT Modelowane prawdopodobieństwo Z/P = 1

Efekt	Kolumna	Ocena	Standard Błąd	Walda Stat.	GU górna 95,0%	GU dolna 95,0%	P	Iloraz szans	Ufność OR -95%	Ufność OR 95%
wyraz wolny	1	-13,500	3,503	14,853	-20,366	-6,635	0,000	0,000	0,000	0,001
Punkty	2	0,155	0,038	16,801	0,081	0,229	0,000	1,168	1,084	1,258
rabp: 8	3	-2,386	0,985	5,874	-4,316	-0,456	0,015	0,092	0,013	0,634
skuteczne 21-24 sek.	4	0,788	0,267	8,699	0,264	1,312	0,003	2,199	1,303	3,713
nieskuteczne 17-20 sek.	5	-0,325	0,132	6,069	-0,584	-0,067	0,014	0,722	0,558	0,936
rzędowe akcje bez pkt: 2	6	0,503	0,184	7,464	0,142	0,863	0,006	1,653	1,153	2,371

W wielowymiarowym modelu regresji logistycznej znalazło się 5 zmiennych diagnostycznych. Przy wzroście zmiennej „21-24 sek.” o jednostkę, następuje ponad dwukrotny wzrost prawdopodobieństwa zwycięstwa, tak samo w przypadku wzrostu wyniku „Punkty” o jednostkę, następuje wzrost prawdopodobieństwa zwycięstwa o 16,8% i analogicznie przy wzroście o jednostkę wyniku zmiennej „Rzędowe akcje bez pkt: 2” następuje wzrost prawdopodobieństwa zwycięstwa o 65,3%. Wzrost o jednostkę zmiennej „rabp: 8” zmniejsza prawdopodobieństwo zwycięstwa prawie ponad jedenastokrotnie (10,87). Wzrost zmiennej „nieskuteczne 17-20 sek.” o jednostkę, zmniejsza prawdopodobieństwo zwycięstwa o 38,4%. Analiza dobroci dopasowania modelu do danych w postaci testu Hosmera Lemeshowa =

12,22 poziom $p= 0,14$ oraz pola powierzchni pod krzywą $AUC=0,90\pm 0,03$ pozwala na stwierdzenie, że nie ma podstaw do stwierdzenia, że model jest źle dopasowany, co potwierdza również poniższa interpretacja w postaci krzywej ROC.



Rysunek 3. Krzywa ROC

5. WNIOSKI

Badania przeprowadzone w niniejszej dysertacji oraz uzyskane wyniki dotyczące czynników determinujących wyniki sportowe, obejmowały zmagania ośmiu najlepszych zespołów w rozegranych Mistrzostwach Świata w koszykówce mężczyzn w Chinach w 2019 roku. Analiza sześćdziesięciu trzech meczów mistrzostw oraz dostarczone w niej dane, pozwoliły na sformułowanie następujących wniosków:

1. Poprzez użycie autorskiego programu do rejestracji i analizy gry w koszykówkę wyodrębniono zbiory nowych elementów diagnostycznych gry. Pierwszy zakres obejmował trzydzieści dwie, znane w literaturze zmienne. Drugi, innowacyjny sektor dotyczył tempa gry poszczególnych zespołów, w swojej liczbie zawierał łącznie,

trzydzieści cztery elementy gry. Trzeci moduł zastosowany po raz pierwszy w pracach naukowych dotyczył akcji rzędowych i składał się z czterdziestu czterech wskaźników.

2. Spośród badanych 110 zmiennych diagnostycznych zredukowano liczbę czynników z materiału badawczego, z czego wyselekcjonowano dwadzieścia wskaźników diagnostycznych. Ostateczny model analizy dyskryminacyjnej przedstawił informacje, dotyczące wyodrębnionych zmiennych, które wyrażały się największym wpływem na końcowy ranking mistrzostw świata. Możliwe stało się utworzenie modelu poprawności klasyfikacji opartej o zmienne diagnostyczne względem zajmowanych miejsc w Mistrzostwach. Poprawność klasyfikacji na podstawie zmiennych diagnostycznych dla czterech grup wyniosła 57,61%, natomiast dla dwóch grup utrzymywała się na wysokim poziomie 83,70%.
3. Na podstawie uzyskanych wyników można stwierdzić, że ranking rzeczywisty i teoretyczny są bardzo silnie powiązane ze sobą ($R=0,88$; $p=0,0039$). Analizy wykazały, iż cztery zespoły uplasowały się na właściwych miejscach, dwa zespoły różniły się o jedną lokatę w stosunku do rankingu rzeczywistego oraz dwa zespoły odbiegały we wspomnianych rankingach o dwie lokaty. Analizy potwierdziły stopień dopasowania i zestawienia modelu teoretycznego z rzeczywistym, uzyskanego z pomocą rankingu Mistrzostw Świata.
4. Na bazie regresji logistycznej dla zmiennych zależnych binarnych wskazanych zostało pięć zmiennych diagnostycznych, które najlepiej prognozują prawdopodobieństwo zwycięstwa i porażki. O wzroście szans na zwycięstwo stanowią: wzrost liczby skutecznych akcji w przedziale czasowym 21- 24 sek., wzrost liczby zdobytych punktów, a także spadek liczby dwóch rzędowych akcji bez punktu, liczby ośmiu rzędowych akcji bez zdobytych punktów oraz liczby nieskutecznych ataków w przedziale 17- 20 sek.

Przedstawione wnioski i argumenty potwierdzają, że metoda pomiaru była skuteczna na wszystkich etapach badań, a wszystkie hipotezy zostały zweryfikowane pozytywnie. Ponadto należy stwierdzić, że zastosowanie jednolitej metody analizy gry w koszykówkę oraz wdrożenie opisanego narzędzia diagnostycznego umożliwi trenerom sektora młodzieżowego i zawodowego na prowadzenie zajęć treningowych w sposób rzeczowy i ukierunkowany na drogę prowadzącą do sukcesu w przyszłych Mistrzostwach Świata lub Mistrzostwach Europy.

BIBLIOGRAFIA

MONOGRAFIE

1. Ryguła Igor, *Elementy teorii, metodyki, diagnostyki i optymalizacji treningu sportowego*, Akademia Wychowania Fizycznego, Katowice 2000.
2. Młodak Andrzej, *Analiza taksonomiczna w statystyce regionalnej*, Wydawnictwo DIFIN, Warszawa 2006
3. Sobczyk Mieczysław, *Statystyka*, Wydawnictwo Naukowe PWN, Warszawa 2002.
4. Starzyńska Wacława, *Statystyka praktyczna*, Wydawnictwo Naukowe PWN, Warszawa 2005.
Zuccolotto Paola, Manisera Marica, *Basketball Data Science: With Applications in R*, Chapman & Hall, CRC Data Science Series, CRC Press 2020.

ARTYKUŁY W PUBLIKACJACH ZBIOROWYCH

1. Dembiński Jacek., *Kwantyfikacja potencjału działania zespołowego a potencjał indywidualny na przykładzie koszykówki cz. I. Założenia teoretyczne*, „Edukacja Wychowanie Sport”, 2021.
2. Huciński Tadeusz, Czerlonko Małgorzata, *Ocena efektywności walki sportowej zespołu w koszykówce*, Roczniki naukowe AWF, Gdańsk 2001, s. 27 – 37.
3. Szwarz Andrzej, Lekner Izabela, Lipińska Patrycja, *The efficiency of action of the top level basketball players with regard to the player position*, Rozpraw Naukowe AWF, Wrocław 2013, nr 42, s. 44 – 61.

ARTYKUŁY W CZASOPISMACH NAUKOWYCH

1. José L., Argudo Francisco. M, Alonso José I., *Effect of basketball mass on shot performance among 9–11 year-old male players*, „International Journal of Sports Science & Coaching”, 2012, vol 7, s. 69-79.
2. Hellwig Zdzisław, *Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowanych kadr*, „Przegląd Statystyczny”, 1968, t. 15, nr 4, s. 307 - 327.
3. Hellwing Zdzisław, *Problem optymalnego wyboru predyktant*, „Przegląd Statystyczny”, 1969, nr 4, s. 221-237.
4. Khlifia Riadh, i in., *Kinematic adjustments in the basketball free throw performed with a reduced hoop diameter rim*, „International Journal of Sports Science & Coaching”, 2012, vol 7, s. 371-382.

5. Matulaitis Kestutis, Stonkus Stanislovas, *Įvairaus amžiaus krepšinio komandų žaidimo rezultatyvumo ir metimų įvairovės rodiklių analizė*, „Baltic Journal of Sport and Health Sciences”, 2009, nr 1.
6. Michońska-Stadnik Anna, *Trafność i rzetelność w badaniach glottodydaktycznych*, „Lingwistyka Stosowana” 2011, nr 4, s. 31-40.
7. Mikołajec Kazimierz, Maszczyk Adam, Zając Tomasz, *Game indicators Determining Sports Performance in the NBA*, „Journal of Human Kinetics”, 2013, t. 37, s. 145 – 151.
8. Spearman Charles., *The proof and measurement of association between two things*, „American Journal of Psychology”, 1904 no 15, s. 72–101.
9. Zambová Denisa, Tománek L’ubor, *An efficiency shooting program for youth basketball players*, „Sport Logia”, 2012, nr 8, s. 87-92.

NETOGRAFIA

1. Encyclopedia of Social Science Research Methods/ SAGE Research Methods Online, dostęp dnia: 20.12.2010
2. https://pl.wikipedia.org/wiki/Analiza_dyskryminacyjna/ dostęp dnia: 01.12.2022
3. https://pl.wikipedia.org/wiki/Metoda_Hellwiga/ dostęp dnia: 02.12.2022
4. https://pl.wikipedia.org/wiki/Metoda_Warda/ dostęp dnia: 04.12.2022
5. <https://mathspace.pl/matematyka/receiver-operating-characteristic-krzywa-roc-czyli-ocena-jakosci-klasyfikacji-czesc-7/> dostęp dnia: 07.12.2022
6. https://mfiles.pl/pl/index.php/Metody_taksonomiczne/ dostęp dnia: 04.12.2022
7. <https://nauka.metodolog.pl/glossary/metryka-euklidesowa/> dostęp dnia: 03.12.2022
8. <https://nauka.metodolog.pl/glossary/test-hosmera-lemeshowa/> dostęp dnia: 05.12.2022