

Analiza wpływu uwarunkowań biologicznych na grę w NBA

Marcin Tkocz, 266547

5 czerwca 2023

Spis treści

1 Cel	2
2 Analizowane dane	2
2.1 Opis danych	2
2.2 Przygotowanie danych	3
3 Analiza danych	5
3.1 Zależność gry koszykarzy od wieku	5
3.1.1 Zależność liczby meczy i minut od wieku	5
3.1.2 Zależność liczby punktów/zbiórek/asyst od wieku	7
3.1.3 Zależność liczby prób rzutów od wieku	10
3.1.4 Zależność skuteczności rzutów od wieku	11
3.1.5 W jakim momencie gracze rozgrywają ostatni sezon w NBA?	12
3.2 Zależność gry koszykarzy od wzrostu	14
3.2.1 Wprowadzanie	14
3.2.2 Zależność liczby minut od wzrostu	15
3.2.3 Zależność liczby punktów/asyst/zbiórek/bloków od wzrostu	15
3.2.4 Zależność skuteczności rzutów od wzrostu	16
3.2.5 Klasyfikacja zawodników na podstawie wzrostu	17
4 Wnioski	20

1 Cel

Celem analizy jest odkrycie zależności między uwarunkowaniami biologicznymi koszykarzy a statystykami dotyczącymi ich gry w NBA. Zweryfikowane też zostaną hipotezy, które zostaną postawione w dalszej części.

Do warunków biologicznych można zaliczyć kilka czynników, takich jak wiek, wzrost, waga, czy pochodzenie. Tutaj zostaną wzięte pod uwagę dwa pierwsze, a więc wiek i wzrost. Można spodziewać się, że wpływają one na grę zawodnika w sposób dużo bardziej odczuwalny niż waga czy wzrost. Co więcej, waga zawodnika nie mówi nic o jego dyspozycji fizycznej (nie wiadomo, czy zawodnik jest bardzo masywny czy bardzo otyły). Z tego powodu w analizie zostaną rozpatrzone jedynie **wiek i wzrost** zawodników.

2 Analizowane dane

2.1 Opis danych

Przedmiotem analizy będą statystyki z **23 kompletnych sezonów** w XXI wieku, począwszy od sezonu 1999-2000 aż do sezonu 2021-2022. W tym czasie sezon NBA rozegrało łącznie **2252 zawodników**.

Dla każdego sezonu dostępne są cztery rodzaje danych, które można w swobodny sposób dobierać do analizowanego czynnika:

1. dane osobiste zawodników (wiek, wzrost, waga...)
2. statystyki łączne w sezonie
3. statystyki średnie na mecz
4. statystyki średnie na minutę

Dane zostały zscrapowane przy użyciu biblioteki BeautifulSoup ze strony <https://basketball.realm.com>.

	Bios	Totals	Per game	Per minute
2000	439	439	439	439
2001	441	441	441	441
2002	440	440	440	440
2003	428	428	428	428
2004	442	442	442	442
2005	464	464	464	464
2006	458	458	458	458
2007	458	458	458	458
2008	451	451	451	451
2009	445	445	445	445
2010	442	442	442	442
2011	452	452	452	452
2012	478	478	478	478
2013	469	469	469	469
2014	482	482	482	482
2015	492	492	492	492
2016	476	476	476	476
2017	486	486	486	486
2018	540	540	540	540
2019	530	530	530	530
2020	529	529	529	529
2021	540	540	540	540
2022	605	605	605	605

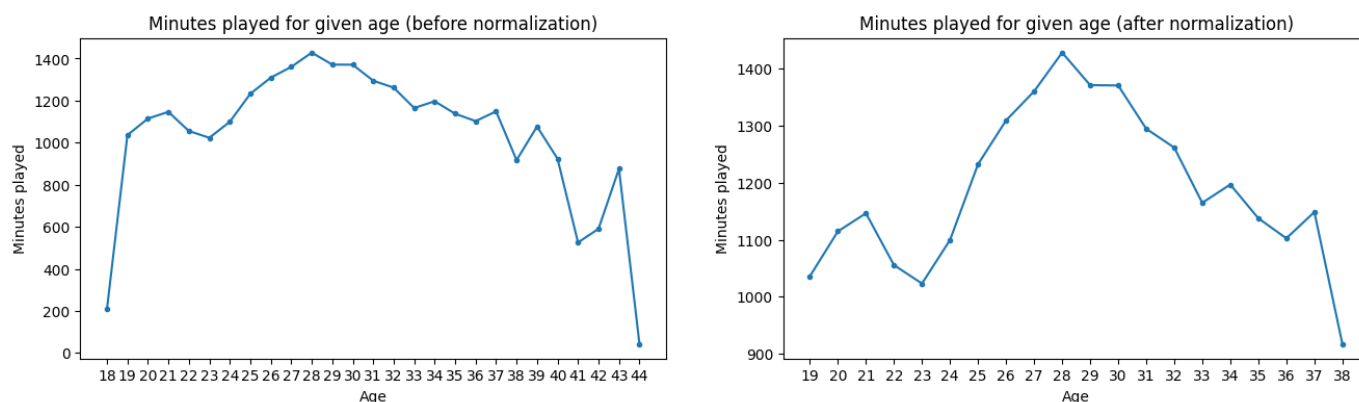
Tabela 1: Liczba koszykarzy dla każdego sezonu i każdej ramki danych

2.2 Przygotowanie danych

Pobrane dane wyrażały wzrost zawodników w systemie amerykańskim (stopy, cale). Przed analizą *wzrost zawodników został przekonwertowany do jednostki centymetrowej*, z której powszechnie korzysta się w Europie.

Player	Height	Player	Height
Steven Adams	6-11	Steven Adams	210.8
Bam Adebayo	6-10	Bam Adebayo	208.3
LaMarcus Aldridge	6-11	LaMarcus Aldridge	210.8
Kyle Alexander	6-10	Kyle Alexander	208.3
Nickeil Alexander-Walker	6-5	Nickeil Alexander-Walker	195.6

Rysunek 1: Wzrost w stopach i centymetrach



Rysunek 2: Przykładowy wykres przed i po normalizacji

Przy analizie danych w większości przypadków będzie wykorzystywana funkcja normalizująca w celu zwiększenia wiarygodności wyników. Rozpatrując daną kategorię (wiek/wzrost), funkcja usuwa te zakresy, dla których nie ma wystarczająco dużej liczby koszykarzy. Dolną granicą w dalszej analizie będzie 50 zawodników w danej kategorii.

Na lewym wykresie na rysunku 2, korzystającym z danych przed normalizacją, odnotowano gwałtowny wzrost średniej liczby rozegranych minut na sezon w wieku 43 lat. Średnia ta jest porównywalna z liczbą minut dla wieku 38, a więc dla zawodników o 5 lat młodszych. Jest to spowodowane faktem, że tylko jeden zawodnik w historii NBA rozegrał sezon w tym wieku - był to Vince Carter, który w sezonie 2019-20, w wieku 43 lat, rozegrał ponad 875 minut w 60 meczach.

Funkcja normalizująca odrzuca kolumny o zbyt małej liczbie danych. Tak przerobiona ramka danych jest przedstawiona na prawym wykresie - przedstawione są statystyki tylko dla tych wieków zawodników, w jakich przynajmniej 50 zawodników w historii NBA rozegrało sezon.

Oczywiście wykres nie pokazuje jeszcze w pełni wyraźnie, że z upływem lat zawodnik rozgrywa coraz mniej minut - wciąż można zaobserwować wzrost tej liczby w wieku 34 i 37 lat, mimo normalizacji. Jest to spowodowane tym, że najstarsi zawodnicy w lidze są często wartościowi dla zespołu mimo obniżonych zdolności fizycznych (w przeciwnym wypadku byłoby już prawdopodobnie na

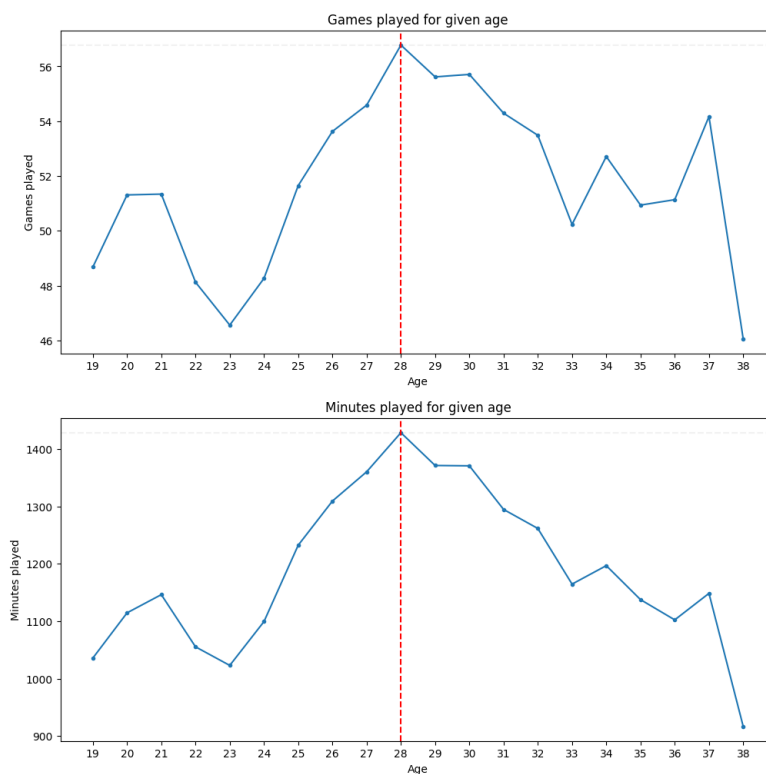
emeryturze). Stąd mimo że fizyczna wydajność zawodników spada z wiekiem, zawodnicy tacy jak Dirk Nowitzki, Karl Malone, Michael Jordan, czy Tim Duncan wciąż byli w stanie osiągnąć ponad 60 meczy w sezonie w wieku 39 lat (nie bez powodu każdy z wymienionych był nagradzany nagrodą MVP i był ponad 10 razy wybierany do NBA All-Star)

3 Analiza danych

3.1 Zależność gry koszykarzy od wieku

3.1.1 Zależność liczby meczy i minut od wieku

Pierwszą hipotezą związaną z wiekiem jest jego silna korelacja z dyspozycją fizyczną zawodników. W większości dyscyplin około 30 roku życia zawodnicy przeżywają szczyt swojej formy, która do końca kariery spada. Można się więc spodziewać, że do około 30 roku życia średnia liczba meczy w sezonie i minut na mecz będzie rosła, a po osiągnięciu maksimum z roku na rok będzie maleć aż do zakończenia kariery.

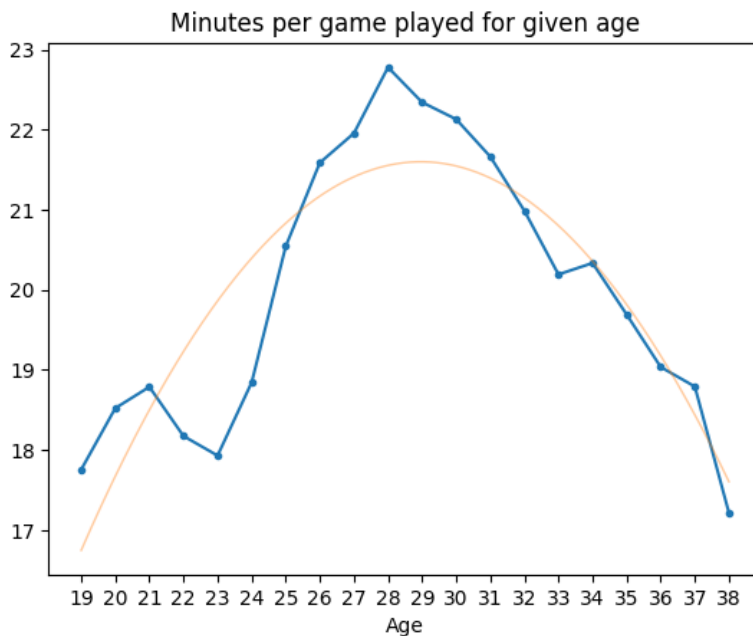


Rysunek 3: Zależność wieku od liczby rozgrywanych meczy i minut na sezon

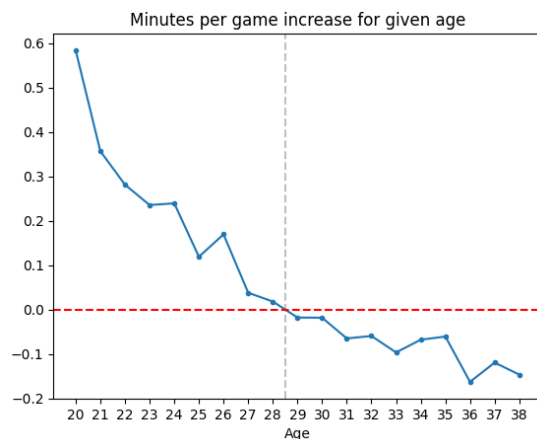
Na rysunku 3 widać, że średnia liczba meczy i minut na sezon rzeczywiście rośnie od 23 roku życia i zaczyna spadać po **28 roku życia, w którym następuje szczyt fizycznych możliwości zawodnika**.

Można również zaobserwować, że statystyki rosną od 19 do 21 roku życia, po czym maleją do 23 roku życia. Tę zależność można zaobserwować zarówno dla liczby meczy i liczby minut na sezon. Prawdopodobnie jest to spowodowane faktem, że najmłodsi zawodnicy NBA (19-21 lat) to gwiazdy amerykańskich uczelni, które grają w większości meczy, choć nieco mniej niż zaprawieni zawodnicy (okres przystosowywania do fizycznych wyzwań NBA). Dla wieków 21-23 statystyki maleją, czego powodem może być przybywanie do NBA młodszych graczy, którzy nie grają tak dobrze jak gwiazdy wzięte do ligi w wieku 19 lat. Obniżony średni poziom gry skutkuje obniżeniem średniej liczby meczy i minut w sezonie. Innym powodem mogłyby być urazy młodych zawodników, którzy będąc nieprzystosowani do fizycznej gry w NBA po grze na poziomie uczelnianym, mogliby nabawić się urazów.

Inną obserwacją jest fakt, że *średnia liczba minut na sezon jest dużo podatniejsza na starzenie* - od 28 do 37 roku życia maleje niemal liniowo. Tego samego nie można powiedzieć o liczbie meczy, co nie powinno być dużym zaskoczeniem - starszy zawodnik może wejść na boisko na ledwie parę minut (np. z szacunku, wymogów kontraktowych), a mecz wciąż zostanie dodany do tej statystyki. Tę zależność potwierdza rysunek 4, ukazujący regresję dla wieku i średniej liczby minut rozgrywanych *na mecz*.



Rysunek 4: Regresja dla wieku i liczby rozgrywanych minut na mecz



Rysunek 5: Spadek przyrostu liczby minut na mecz od wieku

Widać wyraźnie, że średnia liczba rozgrywanych minut na mecz wzrasta do 28 roku życia, a następnie maleje. Wykres jest dość "gładki" od 23 do 28 roku życia, gdy rośnie zarówno liczba meczy i liczba minut w sezonie. Po prawej stronie wykres jest nieco mniej wyrównany, czego powodem jest niejednostajnie malejąca liczba rozgrywanych meczy w sezonie, omówiona wcześniej.

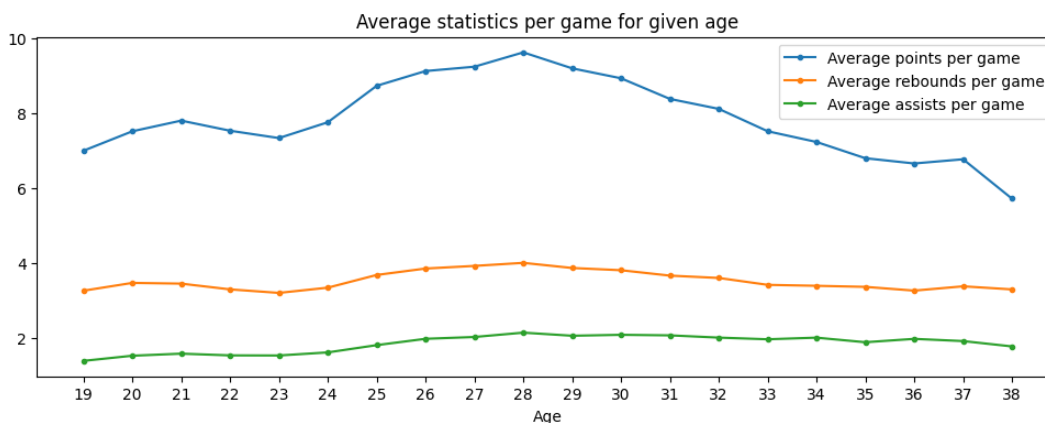
Aby zależność była jeszcze bardziej oczywista, na wykresie na rysunku 5 przedstawiono średnie zmiany w liczbie rozgrywanych minut w stosunku do statystyk z poprzedniego roku (innymi słowy, jeśli $f(x)$ to liczba minut na mecz w wieku x lat, to przyrost jest dany wzorem $g(x) = f(x) / f(x-1)$). Następnie dla każdego wieku obliczono średnią zmianę liczby rozgrywanych minut. Na koniec od przyrostów odjęto 1, aby spadki (wartości stosunkowe mniejsze od 1) były wyrażane przez liczbę ujemną.

Ponownie widać, że do około 28 roku życia średnia liczba rozgrywanych minut na mecz rośnie (wartości większe od zera), jednak tempo tego przyrostu maleje (wartości zbliżają się do zera). Natomiast **średnia zmiana po osiągnięciu 29 jest zawsze ujemna, co oznacza, że przeciętny zawodnik z roku na rok rozgrywa coraz mniej minut na mecz**. Także na tym wykresie można zaobserwować wzrost po 33 roku życia (próba wyjaśnienia pojawi się później), jednak ta interpretacja jest oczywista - rozpatrując wszystkich zawodników w danym wieku średnia liczba minut w sezonie zachowuje się czasem niejednoznacznie, ale gdy rozpatrywani są indywidualni zawodnicy, to po 29 roku życia rozgrywają oni z sezonu na sezon coraz mniej minut.

3.1.2 Zależność liczby punktów/zbiórek/asyst od wieku

Będąc na parkiecie, koszykarz może wspomagać swój zespół w kilku aspektach, z których najważniejsze to: punkty, zbiórki i asysty. Mimo że różni koszykarze specjalizują się w innych aspektach gry, w XXI wieku od niemal każdego zawod-

nika oczekuje się, że będzie zdobywał punkty. Innymi słowy, nie każda statystyka jest sobie równa. Pokazuje to wykres na rysunku 6.



Rysunek 6: Zależność średnich statystyk na mecz od wieku

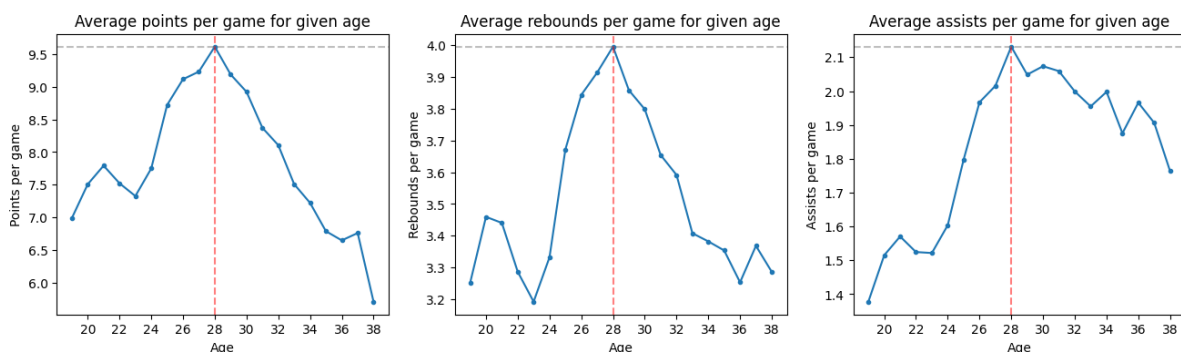
Bez względu na wiek, średnia liczba punktów na mecz zawodników jest ponad dwa razy większa niż liczba zbiórek na mecz, która jest około dwa razy większa od liczby asyst. Punkty nie powinny być zaskoczeniem, ponieważ ostatecznie wygrywa drużyna, która zdobędzie ich więcej, stąd też zawodnicy priorytetyzują zdobywanie punktów. Na drugim miejscu umiejscowiły się zbiórki, które również odgrywają w grze znaczącą rolę - zawodnicy próbują złapać piłkę z powietrza, albo żeby kontynuować atak, albo w celu przejścia z obrony do ataku. Głównie w tym celu każda drużyna koszykarska ma w zespole *centra*, czyli najwyższego zawodnika, którego głównym zadaniem jest zbieranie piłek z powietrza. Na trzecim miejscu plasują się asysty - dogrywanie piłek do niekrytych partnerów ułatwia zdobywanie punktów, ale nowoczesne NBA w dużej mierze opiera się na indywidualnych akcjach zawodników, którzy sami mogą okiwać obrońcę i wykonać rzut. W porównaniu do np. piłki nożnej dużo łatwiej jest też przebiec przez całe boisko bez podawania piłki, stąd sumaryczna liczba asyst jest mniejsza.

Ze względu na skalę, na wykresie 6 nie widać, w jakim wieku przeciętny zawodnik osiąga największą liczbę punktów, zbiórek i asyst. Te wartości zostały przedstawione na rysunku 7.

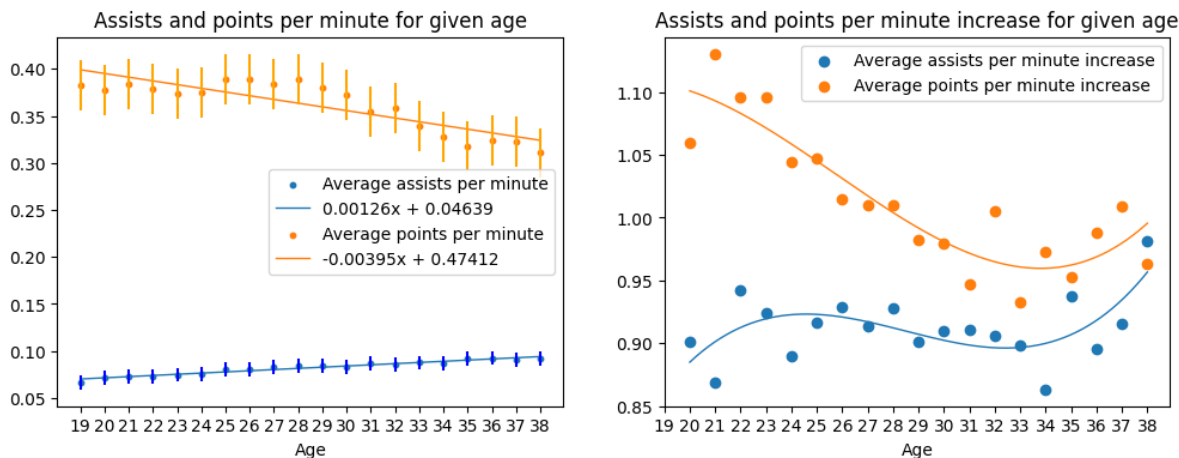
Ponownie wiek 28 lat okazuje się być wiekiem, w którym przeciętny zawodnik osiąga najwyższe osiągi w punktach, zbiórkach i asystach na mecz. Biorąc pod uwagę wcześniejsze wykresy można jednoznacznie stwierdzić, że *w wieku 28 lat przeciętny koszykarz NBA rozgrywa najlepszy sezon, z największą liczbą meczów, punktów, zbiórek i asyst w sezonie.*

Można też sprawdzić jak zmieniają się te statystyki z roku na rok, a więc w jakim stopniu wiek wpływa na osiągnięte wyniki.

Lewy wykres na rysunku 8 pokazuje, że podczas gdy w perspektywie całej kariery średnio zawodnik zdobywa coraz mniej punktów na minutę, liczba asyst



Rysunek 7: Zależność średnich statystyk na mecz od wieku



Rysunek 8: Regresja skuteczności rzutów od wieku

nie tylko nie maleje, ale nawet rośnie. Można więc stwierdzić, że starzenie nie ma tak destruktywnego wpływu na asystowanie jak na zdobywanie punktów. Mogłoby to również sugerować, że z racji coraz gorszych możliwości zdobywania punktów, gracz próbuje wspierać drużynę przez asystowanie.

Z kolei prawy wykres pokazuje średnie zmiany liczby punktów asyst zawodników w porównaniu do poprzedniego roku. Ponownie widać, że przyrost liczby punktów na minutę maleje dość gwałtownie, podczas gdy przyrost liczby asyst na minutę prawie niezależny od wieku (choć mniejszy od 1). Wskazuje to jednoznacznie na fakt, że asysty są znacznie mniej zależne od wieku zawodnika niż zdobywane punkty.

Kolejna hipoteza do sprawdzenia stwierdza, że wiek wpływa na więcej niż tylko liczba punktów - zmianie może ulegać także procentowa skuteczność rzutów (zawodnik doskonali formę/starzenie negatywnie wpływa na formę) oraz

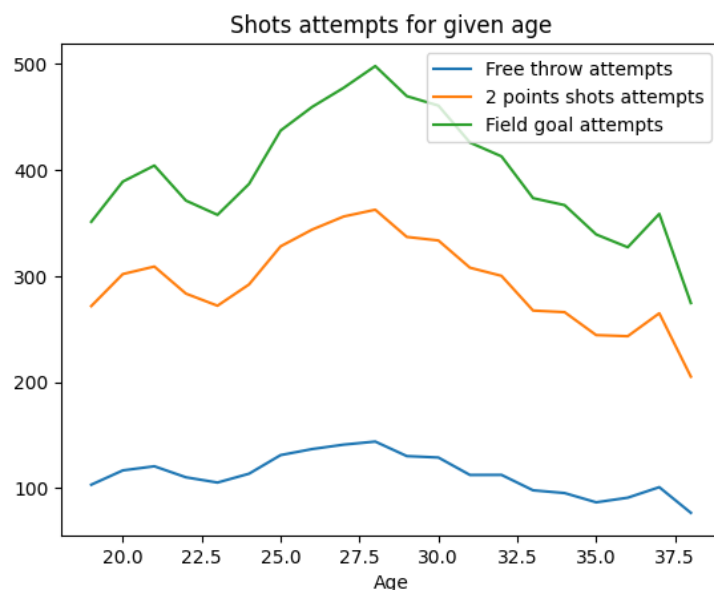
sumaryczna liczba prób zdobycia punktów (więcej prób aby mimo starzenia utrzymać stałą liczbę punktów/mniej prób spowodowane malejącą skutecznością rzutów).

3.1.3 Zależność liczby prób rzutów od wieku

Rzuty można podzielić na rzuty osobiste i rzuty z gry (field goals), które dalej można podzielić na rzuty za dwa punkty i rzuty za trzy punkty.

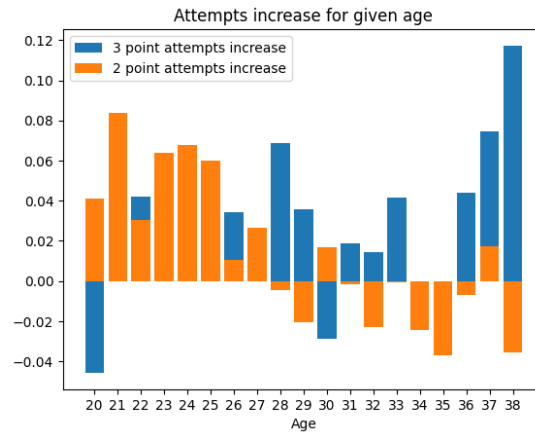
Rzuty za dwa punkty są oddawane z mniejszej odległości, teoretycznie więc łatwiej jest trafić. Z drugiej strony im bliżej kosza stoi zawodnik, tym aktywniej będzie on kryty przez obrońcę z drużyny przeciwnej.

Rzuty za trzy punkty są oddawane zza linii trzech punktów - odległość jest większa, więc trudniej jest trafić. Tym samym zawodnicy mają większą swobodę oddawania rzutów.



Rysunek 9: Zależność średniej liczby rzutów osobistych i z gry w sezonie od wieku

Widać, że wykres 9 prób rzutów przypomina zależności punktów od wieku - nie powinno to dziwić, bo im więcej zawodnik zdobywa punktów, tym więcej musiał wykonać rzutów. Wszystkie liczby prób maleją też od ok. 28 roku życia. Co do liczby rzutów osobistych, są one bezpośrednio skorelowane z liczbą rzutów z gry. Rzuty osobiste są oddawane, gdy zawodnik został sfaulowany podczas wykonywania rzutu. Można się więc spodziewać, że na liczbę wykonywanych rzutów osobistych wpływa czas przebywania zawodnika na boisku oraz miejsce tego przebywania.



Rysunek 10: Średni przyrost prób rzutów w porównaniu do poprzedniego roku

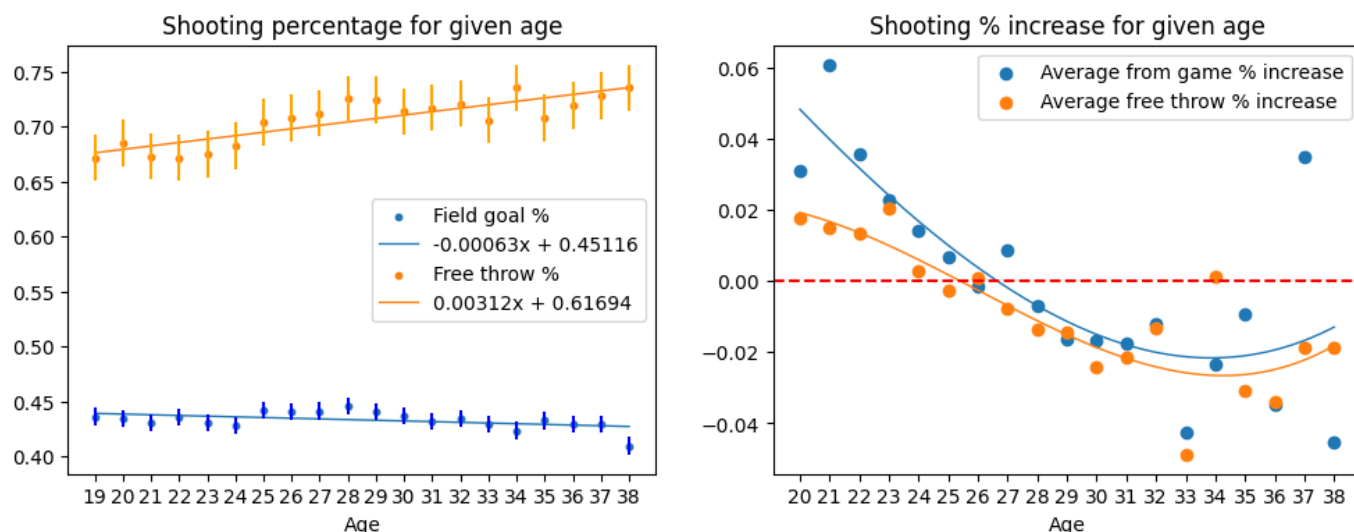
Na wykresie 9 widać, że liczba oddawanych rzutów osobistych rzeczywiście jest skorelowana z liczbą oddawanych rzutów z gry - liczby rosną do około 21 roku życia, maleją do 22 po czym rosną aż do około 28 życia, po czym maleją niemal do końca kariery. Oczywiście wykres liczby rzutów osobistych jest spłaszczony w porównaniu do wykresów rzutów z gry i za 2 punkty, gdyż rzuty osobiste wykonywane są tylko po faulach na zawodnikach wykonujących rzut z gry.

Poprzednią hipotezę co do liczby rzutów można poszerzyć: z upływem lat zawodnicy będą oddawać coraz mniej rzutów za dwa punkty (większa rywalizacja pod koszem, fizyczna gra), a więcej rzutów za trzy punkty, które są więcej punktowane, a przy których odczuwa się słabsze krycie. Zależność między średnim wzrostem liczby rzutów za 2 i za 3 punkty przedstawia rysunek 10.

Wyraźnie widać, że liczba rzutów za 2 punkty spada sezonu na sezon po osiągnięciu około 30 roku życia. Zupełnie odwrotnie jest z rzutami za 3 punkty - średnio po 30 roku życia w żadnym sezonie liczba tych rzutów nie maleje. Co więcej, od 36 roku życia obserwuje się gwałtowny wzrost. To sugeruje, że starzejący się zawodnicy rzeczywiście coraz bardziej preferują oddawanie rzutów z mniej strzeżonych pozycji, za które mogą zdobyć więcej punktów, niż za rzuty spod kosza. Nie oznacza to oczywiście, że oddają więcej rzutów za 3 punkty niż za 2 punkty - oddają natomiast stosunkowo coraz więcej rzutów za trzy punkty i coraz mniej za dwa, niż w poprzednich latach kariery.

3.1.4 Zależność skuteczności rzutów od wieku

Na lewym wykresie na rysunku 11 wyraźnie widać, że średnia skuteczność rzutów osobistych w perspektywie 18-letniej kariery jest rosnąca. Dla rzutów z gry przeciwnie - skuteczność w perspektywie kariery maleje w stałym tempie, stąd mniejsze odchylenie standardowe. Jest to najprawdopodobniej spowodowane



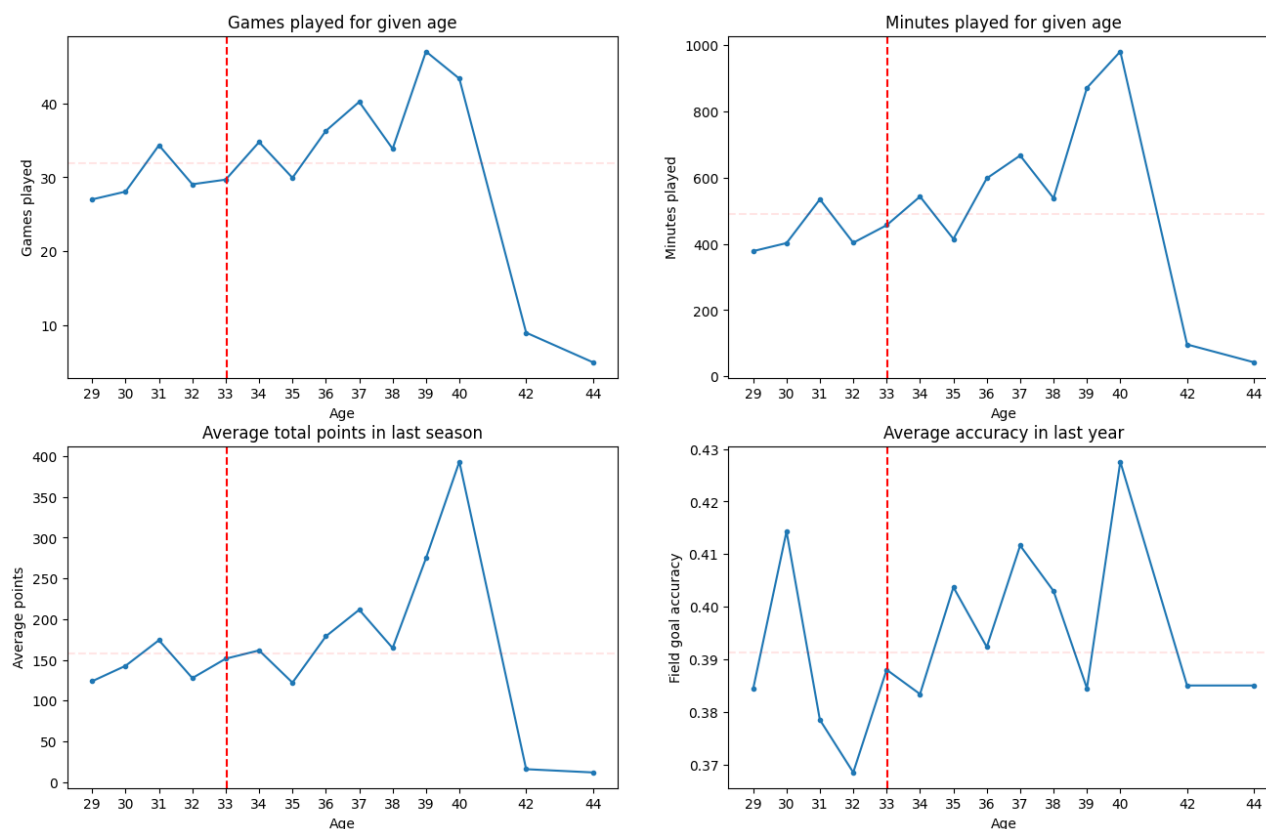
Rysunek 11: Regresja skuteczności rzutów od wieku

wane faktem, że na rzut osobisty ma wpływ tylko i wyłącznie gracz oddający rzut - nikt nie zasłania kosza ani nie może przeszkadzać, stąd pole do doskonalenia rzutów osobistych jest dużo większe niż dla rzutów z gry. Można też zauważyć, że do 24 roku życia skuteczność jest poniżej średniej skuteczności w karierze, następnie do około 30 roku życia skuteczność jest większa niż średnia skuteczność w karierze, by na koniec ponownie spaść poniżej średniej.

Na prawym wykresie widać przeciętny wzrost skuteczności z roku na rok dla zawodników w różnym wieku. Skuteczność rzutów rzeczywiście rośnie do około 25 roku życia. Widać również, że zmiany skuteczności rzutów z pola zmieniają się gwałtowniej niż zmiany skuteczności rzutów osobistych - powodem jest prawdopodobnie to, że na skuteczność rzutów z pola wpływa nie tylko forma rzutu, ale także skuteczność obrony przeciwnika, wysokość wyskoku, oszacowanie zmiennej odległości od kosza itp. Interesujący może być natomiast fakt spłaszczenia krzywej spadkowej w ostatnich latach gry. Jest to prawdopodobnie spowodowane faktem, że większość zawodników grających w tym wieku to gwiazdy NBA, które mają na koncie liczne osiągnięcia i mimo wieku są w bardzo wysokiej formie. Procentowy udział jakościowych zawodników w starszym wieku jest dużo wyższy niż w latach szczytu fizycznych możliwości, stąd średnia skuteczność rzutów maleje w stopniu dużo mniejszym niż u przeciętnych zawodników lub nawet może rosnąć.

3.1.5 W jakim momencie gracze rozgrywają ostatni sezon w NBA?

Ostatnie pytanie które zostanie zadane w kwestii wieku graczy to w jakim wieku zawodnicy rozgrywają przeciętnie swój ostatni sezon w NBA.



Rysunek 12: Statystyki w ostatnim roku gry w NBA

Oczywiście nie każdy ostatni sezon w NBA oznacza zakończenie kariery ze względu na słabe statystyki. Zawodnik może zmienić ligę koszykarską lub zakończyć grę z powodów osobistych/kontuzji. Z dużą pewnością można jednak stwierdzić, że przeważająca większość graczy po 28 roku życia (szczyt kariery) rozgrywa w NBA ostatni sezon, ponieważ kluby nie potrzebują już ich wkładu w grę.

Rozważani będą tylko zawodnicy w wieku powyżej 28 lat, którzy nie grali ani w sezonie 2020-2021, ani w sezonie 2021-2022 (ostatnie dwa analizowane sezony) - eliminujemy tym mylące statystyki zawodników, którzy wciąż mogą grać w obecnym sezonie 2022-2023 (nie zakończyli jeszcze kariery) lub którzy z powodu kontuzji musieli opuścić sezon 2021-2022 a grali w poprzednim sezonie. Dane nie zostały znormalizowane w celu uwzględnienia także starszych koszykarzy. Łączna liczba kwalifikujących się zawodników to 686.

Okazuje się, że średni wiek zawodników z tej grupy to **33 lata** - tym wieku przeciętny zawodnik po 28 roku życia rozgrywa w NBA swój ostatni sezon. Wynik jest o tyle interesujący, że na wykresach przedstawiających średnią liczbę

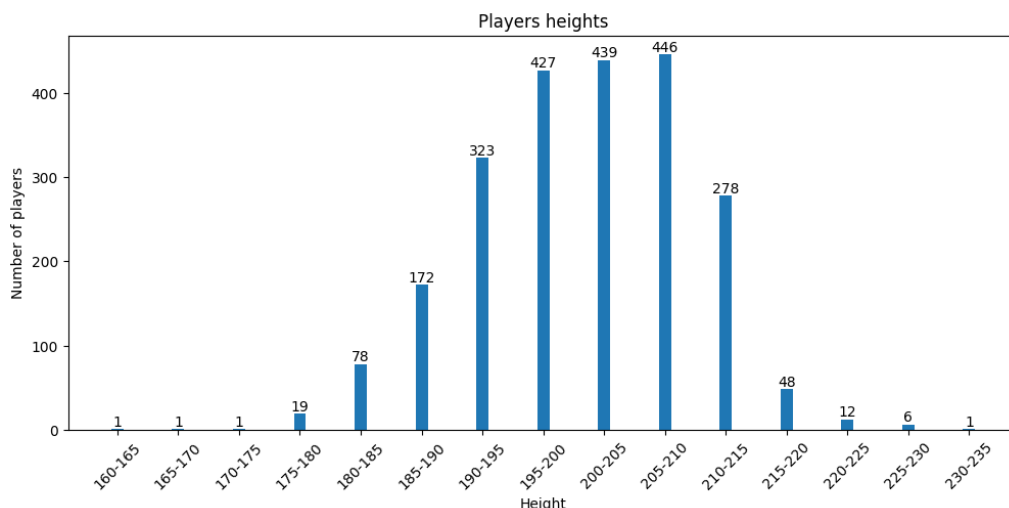
rozgrywanych meczy i minut w sezonie w 33 roku życia odnotowuje się lokalne minimum, a na wykresach dotyczących punktów, zbiórek i asyst obserwuje się gwałtowny spadek średniej. Około 33 roku życia można zauważyć także punkt przegięcia na wykresie obrazującym zmiany skuteczności rzutów osobistych i z gry. Wygląda więc na to, że w tym wieku większość starzejących się zawodników osiąga statystyki skłaniające ich do zakończenia gry w NBA. Po ich odejściu stosunek jakościowych graczy powyżej 33 roku życia zwiększa się, stąd też można zaobserwować w tej grupie wzrost w średniej liczby rozgrywanych meczy i minut na sezon, jak również korzystniejsze statystyki dotyczące asyst i punktów na minutę oraz skuteczności rzutów.

Podobne wnioski można wyciągnąć na podstawie wykresów na rysunku 12, obrazujących średnią liczbę meczy, minut i punktów oraz średnią skuteczność zawodników z tej grupy.

3.2 Zależność gry koszykarzy od wzrostu

3.2.1 Wprowadzanie

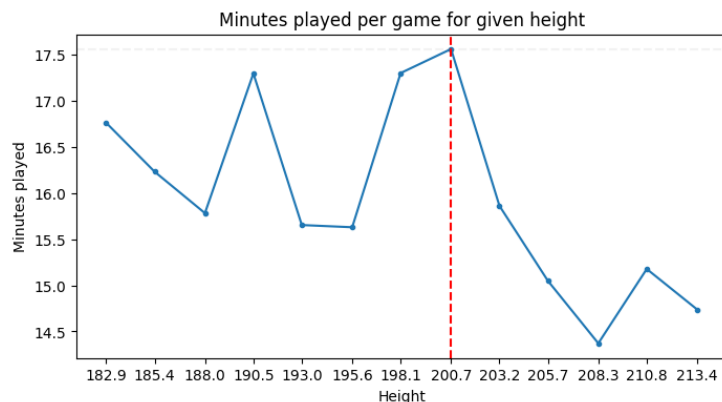
Nie ulega wątpliwości, że wzrost zawodnika ma wpływ na styl jego gry. Najprawdopodobniej daje mu także przewagę nad innymi graczami, skoro tak wielu koszykarzy ma ponad dwa metry. Poniżej zostanie dokonana analiza zależności między wzrostem zawodników a ich statystykami.



Rysunek 13: Rozkład wzrostu zawodników NBA. 93% zawodników mieści się w przedziale 185-215cm

3.2.2 Zależność liczby minut od wzrostu

NBA w XXI wieku wygląda zupełnie inaczej niż parę dekad temu. Dzisiejsza dynamika gry jest dużo większa, zawodnicy muszą biegać od swojego kosza do drugiego w pełnej szybkości. W tym aspekcie wzrost nie wydaje się dawać przewagi. Im większy wzrost, tym z reguły większa masa, więc ciało potrzebuje więcej tlenu podczas wysiłku. Można więc wysunąć hipotezę, że średnia liczba minut na mecz wyższych zawodników będzie mniejsza, niż zawodników o niższym wzroście.

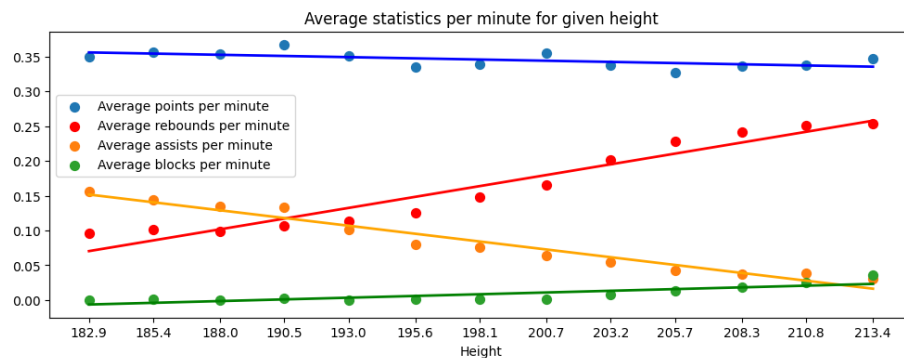


Rysunek 14: Zależność liczby minut na mecz od wzrostu

Rysunek 14 potwierdza te podejrzenia. *Zawodnicy powyżej 200 centymetrów rozgrywają coraz mniej minut na mecz*, co potwierdzałoby postawioną hipotezę. Niejednoznaczny rozkład dla wzrostów poniżej 200 cm może mieć związek z niejasną jeszcze rolą wzrostu - zawodnicy niżsi mogą rozgrywać mniej minut na mecz nie ze względu na zdolności fizyczne, ale na mniejsze możliwości pomocy drużynie związane z niskim wzrostem.

3.2.3 Zależność liczby punktów/asyst/zbiórek/bloków od wzrostu

Tutaj również nasuwa się oczywista hipoteza: im wyższy jest zawodnik, tym łatwiej będzie mu złapać piłkę z powietrza - odnotuje więc więcej zbiórek. Z reguły najwyżsi koszykarze stoją w pobliżu kosza w celu wykonywania krótkich rzutów po podaniach od kolegów - sami z reguły nie podają, będąc w korzystniejszej pozycji do zdobycia punktów. Można się więc spodziewać, że liczba asyst także będzie skorelowana ze wzrostem. Tak samo powinno być z blokami - im zawodnik jest wyższy, tym łatwiej będzie mu zablokować przeciwnika. Trudno natomiast określić, jaka będzie zależność między wzrostem a zdobywanymi punktami. Z jednej strony center ma bliżej do kosza, z drugiej - zdobywa jedynie punkty za 2 i jest kryty przez centra drugiej drużyny. Zależności między wzrostem a statystykami przedstawia rysunek 15.

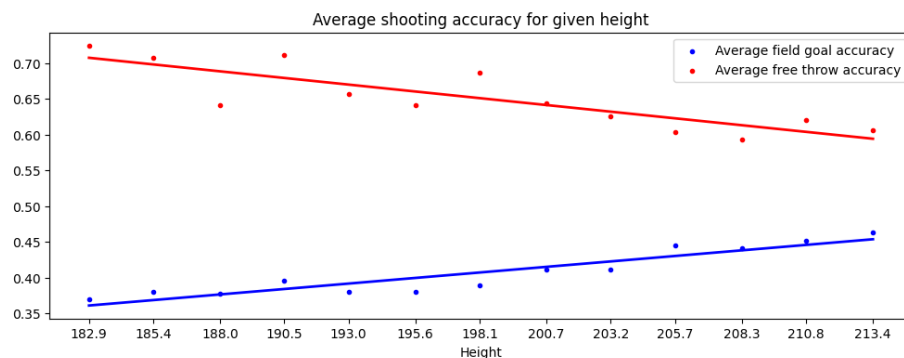


Rysunek 15: Regresja statystyk od wzrostu zawodników

Zgodnie z przewidywaniami, wzrost jest silnie skorelowany z liczbą asyst i zbiórek na mecz - im wyższy jest zawodnik, tym mniej asyst wykonuje, a więcej zbiórek. Rośnie także liczba bloków, które są najrzadszymi ze wymienionych statystyk - najwyżsi zawodnicy wykonują ok. 0.05 bloków na minutę, ponad 5 razy więcej niż najniżsi zawodnicy (w perspektywie sezonu jest to różnica mniej więcej 5 do 30 bloków). Z wykresu nie wynika natomiast, aby istniała jakakolwiek zależność między wzrostem a średnim dorobkiem punktowym - dla wszystkich wzrostów średnia liczba punktów na minutę utrzymuje się na poziomie ok. 0.35, bardzo małe jest także średnie odchylenie standardowe.

3.2.4 Zależność skuteczności rzutów od wzrostu

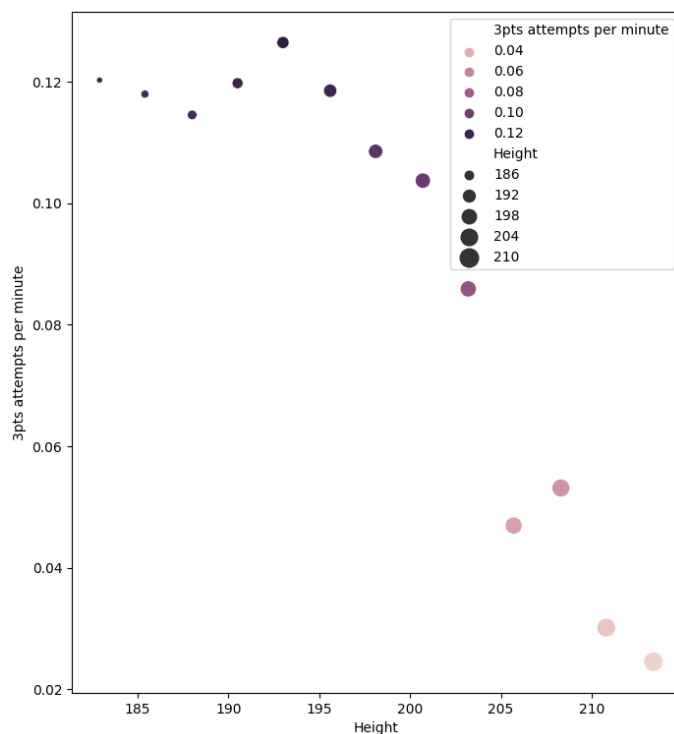
Centrzy stojący pod koszem mają zazwyczaj użytkową technikę rzutu - umieją umieścić piłkę w koszu z bliskiej odległości, jednak dalsze rzuty zaczynają być kłopotliwe, gdyż wymagają odpowiedniej i wytrenowanej formy rzutu, której centrzy zazwyczaj nie mają. Tę hipotezę potwierdza wykres 16.



Rysunek 16: Regresja skuteczności rzutów od wzrostu zawodników

Jak widać, wraz ze wzrostem maleje średnia skuteczność rzutów osobistych, rośnie natomiast skuteczność rzutów z gry, co potwierdza postawioną hipotezę. Można też zauważyć, że odchylenie standardowe jest dużo mniejsze dla skuteczności rzutów z gry, co sugerowałoby, że istnieje między nią ścisła zależność ze wzrostem.

Ostatnia hipoteza jest najbardziej oczywista - im przeciętny zawodnik jest wyższy, tym rzadziej będzie oddawał rzuty za trzy punkty. Zostało już powiedziane, że centry dysponują zazwyczaj użytkową techniką rzutu, przystosowaną do odległości 0-3 metry od kosza. Tacy zawodnicy nie mają więc zazwyczaj żadnego powodu do oddania rzutu za 3 punkty, chyba że kończy się czas na wykonanie akcji i nie ma żadnego innego rozwiązania.



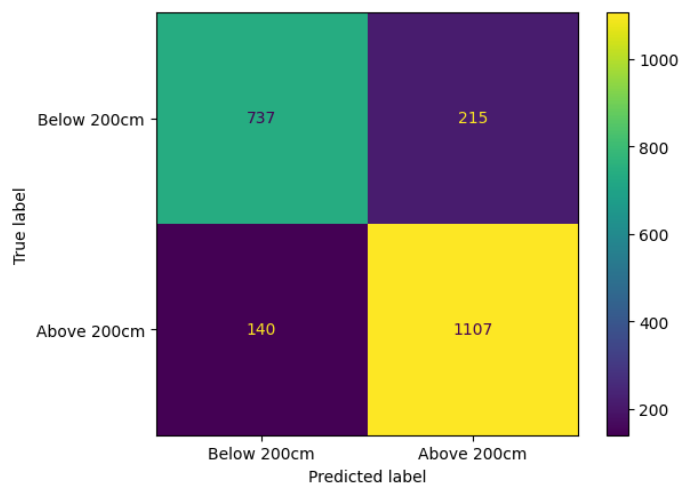
Rysunek 17: Zależność liczby prób rzutów za 3 punkty od wzrostu zawodników

Średnia liczba prób rzutów maleje gwałtownie od ok. 195 centymetrów. Dla zawodników niższych liczba utrzymuje się na mniej więcej stałym poziomie równym 0.12 prób na minutę.

3.2.5 Klasyfikacja zawodników na podstawie wzrostu

Zostało już pokazane, że istnieją zależności między wzrostem zawodników a liczbą zbiórek, asyst i bloków na mecz, skutecznością rzutów osobistych i z pola

oraz liczbą prób rzutów za 3 punkty. Korzystając z kombinacji tych parametrów, z dużą skutecznością można oszacować wzrost zawodnika.



Rysunek 18: Tablica pomyłek klasyfikacji zawodników na niższych i wyższych niż 200cm

Rysunek 18 ukazuje tablicę pomyłek analizowanego modelu dla wcześniej wymienionych parametrów. Podział polegał na określeniu, czy zawodnik ma więcej czy mniej niż 2 metry.

Dokładność klasyfikacji wyniosła 0.839. Można więc stwierdzić, że na podstawie serii przedstawionych parametrów decyzyjnych można z niemal 85% pewnością określić, czy zawodnik ma mniej czy więcej niż 200cm.

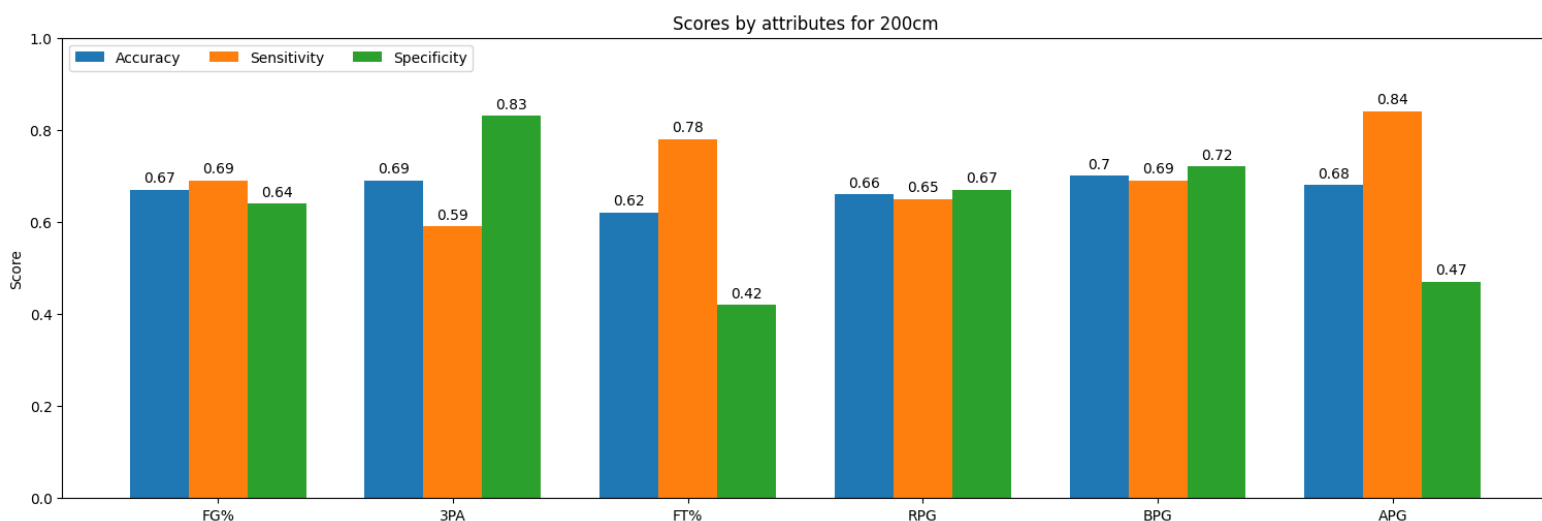
Innymi interesującymi wartościami oceniającymi wyniki testu jest czułość i swoistość.

Czułość określa, z jaką dokładnością model był stanie poprawnie sklasyfikować pozytywnego zawodnika, a więc zawodnika ze wzrostem powyżej 200cm. Z kolei swoistość mówi o dokładności modelu w klasyfikacji zawodników negatywnych, których wzrost nie przekraczał 200 cm.

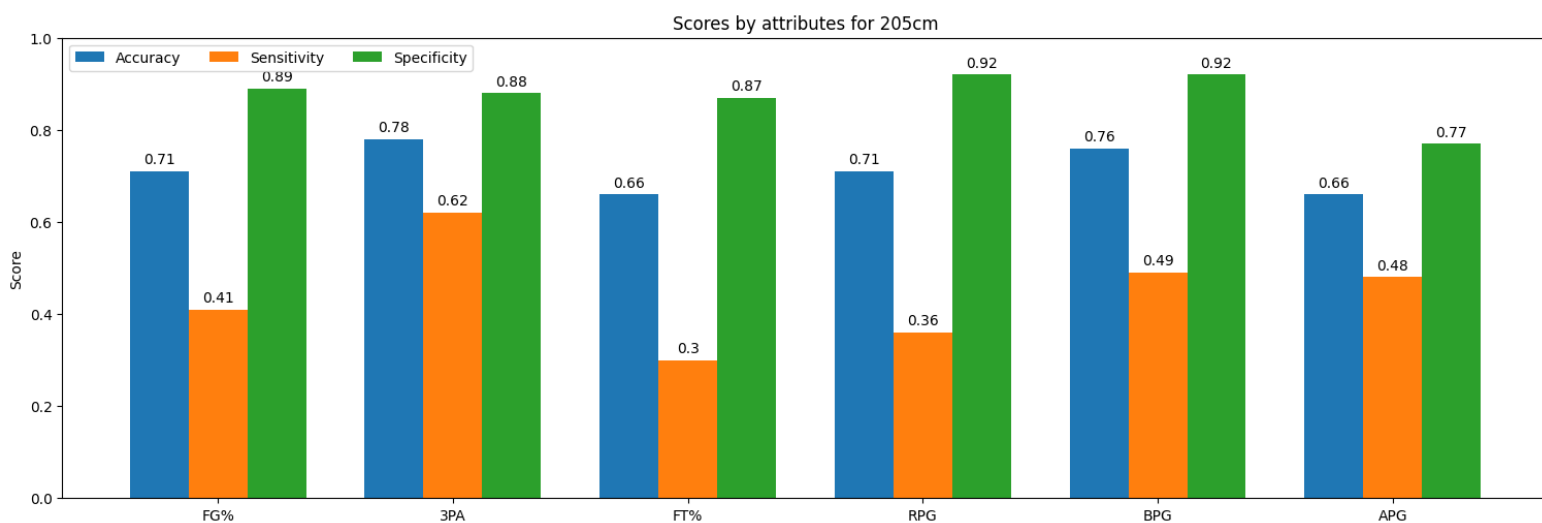
W wynikowej tablicy pomyłek **czułość wyniosła aż 88.77%**. Gdy więc dochodziło do analizy zawodnika mierzącego więcej niż 2 metry, model w niemal 90% był w stanie poprawnie go sklasyfikować. Klasyfikacja zawodników niższych wypadła nieco gorzej, gdyż swoistość modelu wyniosła 77.42%.

Przedstawione wyniki dotyczyły kolektywnej analizy sześciu parametrów. Pozostaje odpowiedzieć na pytanie - który z nich w sposób dominujący cechuje wysokich zawodników oraz jak dokładnie da się określić przedział wzrostu koszykarza na ich podstawie.

Rysunki 19 i 20 obrazują wyniki klasyfikacji zawodników. Wykres 19 dzielił zawodników na niższych i wyższych niż 200cm, a wykres 20 na zawodników niższych i wyższych niż 205cm. W obu przypadkach zawodnicy byli kolejno



Rysunek 19: Wyniki klasyfikacji zawodników na niższych i wyższych niż 200cm na podstawie poszczególnych parametrów



Rysunek 20: Wyniki klasyfikacji zawodników na niższych i wyższych niż 205cm na podstawie poszczególnych parametrów

klasyfikowani na podstawie pojedynczego parametru. Wykresy ukazują wartości dokładności, czułości i swoistości modeli dla danych parametrów.

Pierwszym z wniosków jest fakt, że model klasyfikujący zawodników na wzrost 200cm ma mniejszą średnią dokładność, wynoszącą 0.67 dla pierwszego

modelu i 0.71 dla drugiego. Jest to oczywiście spowodowane faktem, że im wyższa jest granica wzrostu, tym łatwiej jest przydzielić zawodnika do jednej z grup. Dużo wyraźniej widać różnicę między zawodnikami mierzącymi 230cm i niższymi, niż między zawodnikami mierzącymi 200cm i niższymi.

Druga obserwacja dotyczy wartości czułości i swoistości w obu modelach. Średnia swoistość jest dużo wyższa w drugim modelu klasyfikacji, podczas gdy jego średnia czułość jest dużo niższa.

Im wyższa jest granica wzrostu zawodników, tym łatwiej stwierdzić, że zawodnik nie klasyfikuje się jako zawodnik wysoki. Mówi o tym zwiększona średnia swoistość w drugim modelu - zawodnik niemierzący ponad 205cm może być łatwiej wykryty, skoro większa jest różnica parametrów klasyfikujących zawodników mierzących 205cm w porównaniu do zawodników niższych.

Mniejsza średnia czułość oznacza, że trudniej jest sklasyfikować zawodnika mierzącego ponad 205cm niż zawodnika mierzącego ponad 200cm, a więc że różnica między zawodnikami mierzącymi 200 i 205cm nie jest tak wyraźna, jak różnica między zawodnikami mierzącymi ponad 200cm i mniej niż 200cm.

Na podstawie wykresów można więc stwierdzić, że parametrem najdokładniej klasyfikującym zawodników na grupy wzrostowe więcej/mniej niż 200cm jest liczba prób rzutów za trzy punkty (średnia dokładność 0.735) oraz liczba bloków na mecz (średnia dokładność 0.73). Gdy chodzi o czułość wykrywania zawodników wysokich, najlepiej sprawdza się liczba asyst na mecz (średnia czułość 0.66), oraz ponownie liczba prób rzutów za trzy punkty i liczba bloków na mecz (czułości 0.605 i 0.59). Te dwa parametry ponownie okazują się dominujące przy wykrywaniu zawodników niższych, ze średnią swoistością 0.855 dla rzutów za 3 punkty i 0.82 dla liczby bloków.

Im wyższa będzie granica wzrostu tym korzystniejsza będzie osiągnięta dokładność i swoistość modelu. Z kolei czułość modelu będzie malała wraz z przesuwaniem granicy wzrostu do góry.

4 Wnioski

W tym dokumencie dokonano analizy wpływu uwarunkowań biologicznych na grę koszykarzy w NBA.

Najpierw zbadano zależność gry od wieku. Okazało się, że koszykarze w NBA osiągają szczyt kariery w wieku około 28 lat, gdy rozgrywają najwięcej meczy i minut na sezon, ale również odnotowują średnio najwięcej punktów, asyst i zbiórek. Po osiągnięciu tego wieku, średnia liczba minut na mecz maleje z roku na rok aż do ostatniego sezonu, który następuje przeciętnie wieku około 33 lat. Podobnie maleje średnia liczba punktów na minutę, w przeciwieństwie do liczby asyst, która utrzymuje się na tym samym poziomie dla zawodników w każdym wieku, a nawet wydaje się delikatnie rosnać z wiekiem.

Z wiekiem maleje średnia sumaryczna liczba wykonywanych rzutów. Wyjątkiem są rzuty za trzy punkty, których liczba rośnie, w przeciwieństwie do rzutów za dwa punkty. Średnia skuteczność rzutów z gry utrzymuje się na mniej więcej stałym poziomie przez całą karierę, natomiast skuteczność rzutów osobistych w

perspektywie kariery rośnie.

Analiza wpływu wzrostu koszykarzy na ich grę wykazała, że najwyżsi zawodnicy grają średnio mniej minut na mecz, ale zdobywają ponad dwa razy więcej zbiórek i ponad 5 razy więcej bloków. Z większym wzrostem wiąże się także mniejsza liczba asyst, nie wykazano natomiast związku między wzrostem a liczbą zdobywanych punktów.

Skuteczność rzutów z gry rośnie razem ze wzrostem, od ok. 0.37 dla wzrostu 183 cm do ok. 0.45 dla zawodników mierzących ponad 210 cm. Z kolei skuteczność rzutów osobistych jest dużo niższa dla wyższych zawodników. Gwałtownie maleje także liczba rzutów za trzy punkty na minutę.

Korzystając z tych właściwości dokonano klasyfikacji zawodników na dwie grupy wzrostowe z dokładnością ok. 84%, a czułością sięgającą 89%. Z wymienionych parametrów największą dokładność i swoistość klasyfikacji zapewnia liczba rzutów za trzy punkty oraz liczba bloków na mecz, natomiast największą czułość daje liczba asyst na mecz.