

Raport dotyczący statystyk piłkarzy ligi angielskiej

Filip Mordarski

2 12 2020

Spis treści

Wstęp	1
Opis zbioru danych	1
Wstępna analiza statystyk zawodników	2
Bramkarze	4
Analiza skuteczności bronienia własnej bramki	4
Gra nogami bramkarzy	8
Analiza statystyk obrońców	12
Analiza defensywnych statystyk obrońców	13

Wstęp

Celem niniejszego raportu jest analiza statystyk zawodników, występujących w najwyższej klasie męskich rozgrywek piłkarskich w Anglii w sezonie 2019/2020. Może on okazać się bardzo wartościowy dla trenerów, którym pomoże w poprawianiu jakości gry swojego zespołu. Taki raport z pewnością okaże się korzystny podczas budowania kadry pierwszego zespołu na kolejny sezon. Umożliwi wyszukanie zawodników o charakterystyce pasującej do gry danego zespołu.

Opis zbioru danych

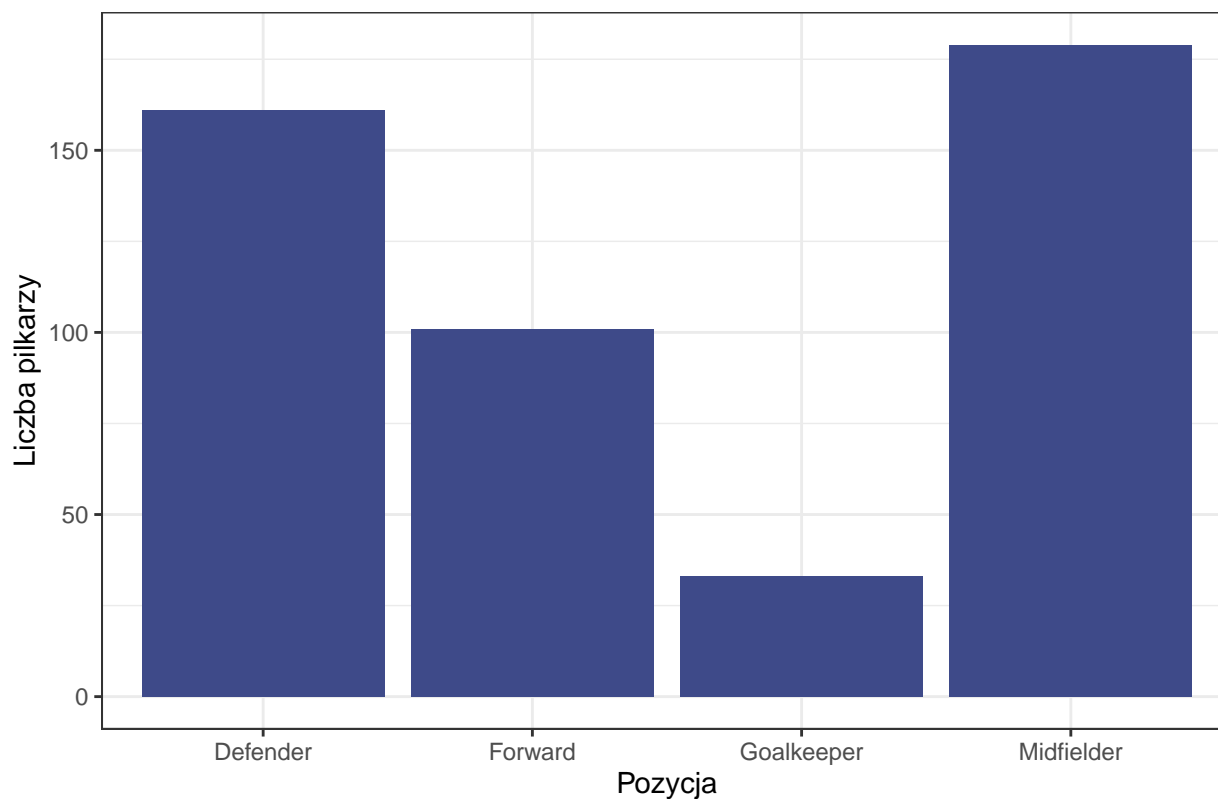
Poniższy rozdział będzie poświęcony opisowi zbioru danych oraz jego statystykom opisowym. Zbiór danych, wykorzystany do przeprowadzenia analizy pochodzi z następującej strony internetowej: <https://www.kaggle.com/krishanthbarkav/english-premier-leaguepl-player-statistics>. Został on utworzony na podstawie statystyk zawodników, występujących w najwyższej klasie męskich rozgrywek piłkarskich w Anglii w sezonie 2019/2020. Zbiór danych zawiera 974 obserwacji oraz 54 zmienne. Jedna obserwacja, odpowiada jednemu zawodnikowi oraz jego statystykom z całego sezonu 2019/2020. Analiza będzie obejmować tylko zawodników, którzy wystąpili przynajmniej raz podczas całego sezonu. Z tego powodu analizowany zbiór danych liczy 474 obserwacji.

Zbiór danych zawiera bardzo dużo zmiennych, aż 54. Przedstawianie statystyk opisowych każdej z nich jest zbędne i nie byłoby wartościowe. Z tego względu zostaną przedstawione jedynie statystyki tych zmiennych, które będą wykorzystane w dalszej analizie.

Wstępna analiza statystyk zawodników

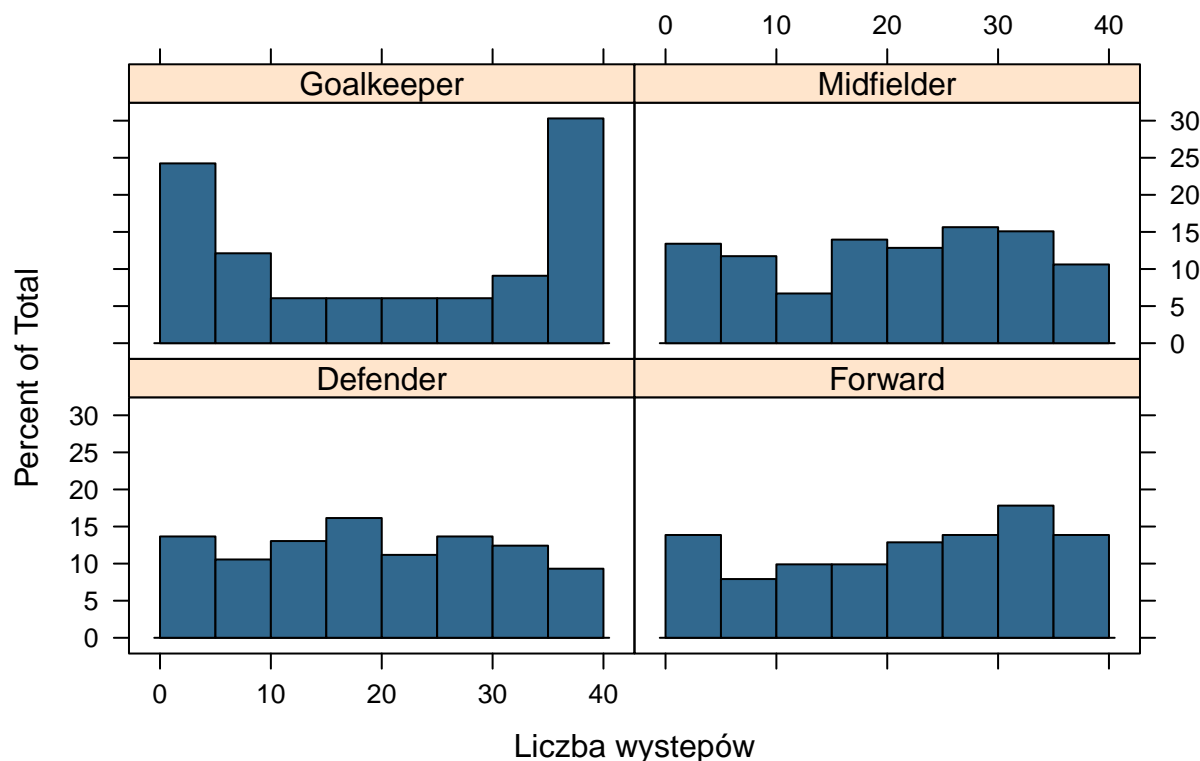
Poniższy rozdział będzie zawierał analizę ogólnych statystyk zawodników z wszystkich pozycji na boisku. W pierwszej kolejności został przedstawiony histogram pozycji zawodników.

Histogram pozycji zawodników



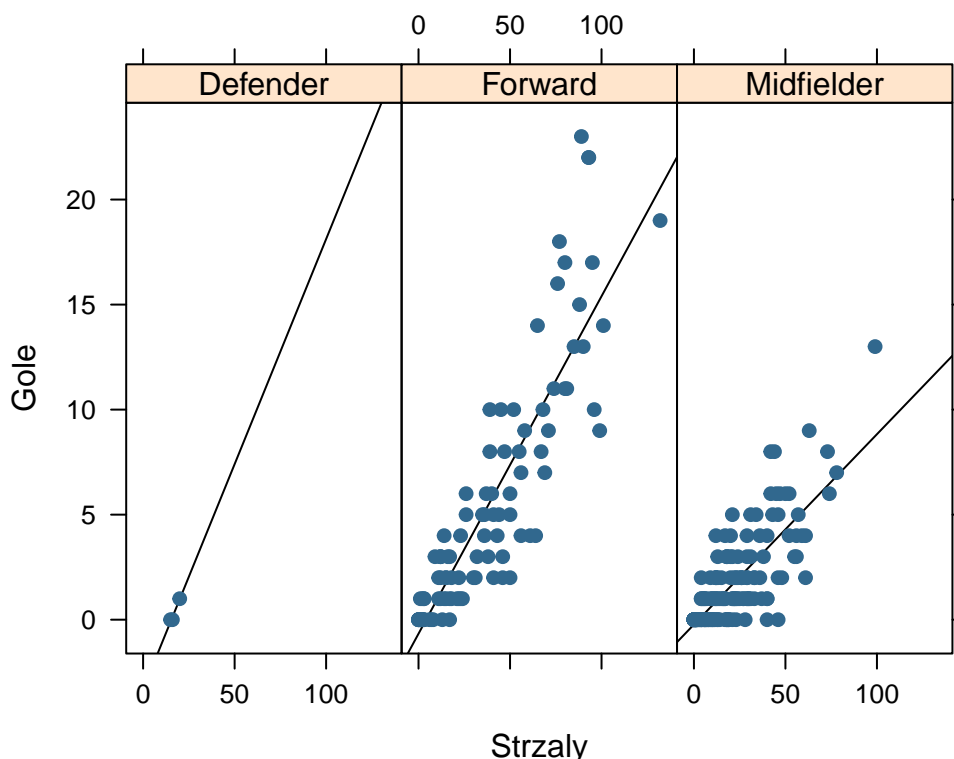
Jak widać na powyższym wykresie, najczęstszą pozycją zawodników jest pomocnik. Porównywalną ilość stanowią obrońcy. Najmniej w lidze angielskiej jest bramkarzy. Jest to dosyć naturalne, ponieważ w meczu w jednej drużynie gra tylko jeden bramkarz. Relatywnie mało jest również napastników. Drużyny często grają w systemie z jednym lub dwoma napastnikami, dlatego ta pozycja nie jest tak mocna obsadzona jak obrona czy pomoc. Następnie został utworzony histogram występów zawodników w podziale na pozycje.

Histogram występów zawodników



Na podstawie powyższego wykresu można zauważyć, że aż 30 % bramkarzy wystąpiło w 35 - 40 meczach. Jest to największy procent wśród pozycji zawodników. Na podstawie tego można wysnuć wniosek, że drużyny rzadziej decydują się zmieniać bramkarzy w podstawowym składzie niż zawodników z innych pozycji. Na innych pozycjach wartości dla danych zakresów występów są podobne. Ciekawy przypadek występuje wśród napastników, którzy zagraли 30 - 35 meczy. Jest ich prawie 20 % wśród innych zawodników na tej pozycji. Również ta wartość może sugerować, że trenerzy raczej nie są skory do zmiany podstawowych napastników w ciągu całego sezonu. Prawdopodobnie najważniejszą statystyką wśród piłkarzy są zdobyte przez nich bramki. To one decydują o ostatecznym wyniku. Przeanalizowano więc zależność pomiędzy zdobytymi bramkami a oddanymi strzałami w zależności od pozycji zawodnika. W tej analizie nie zostaną uwzględnieni bramkarze, ponieważ żaden z nich nie zdobył w tym sezonie bramki.

Zależność goli od strzałów w podziale na pozycje wraz z linią regresji



Powyższy wykres dostarcza wniosków, które dla nawet początkującego fana piłki nożnej wydadzą się dość naturalne i oczywiste. W pierwszej kolejności, występuje pozytywna korelacja pomiędzy ilością bramek a strzałami. Na wykresie można zauważyć, że najczęściej bramek zdobywali napastnicy. Co ciekawe, nachylenie linii regresji jest wyższe niż u pomocników, co sugeruje że wraz ze zwiększeniem liczby strzałów o jeden, ilość bramek wzrasta szybciej u napastników niż u pomocników. Jak można się było również spodziewać, ilość bramek zdobytych przez obrońców jest znikoma.

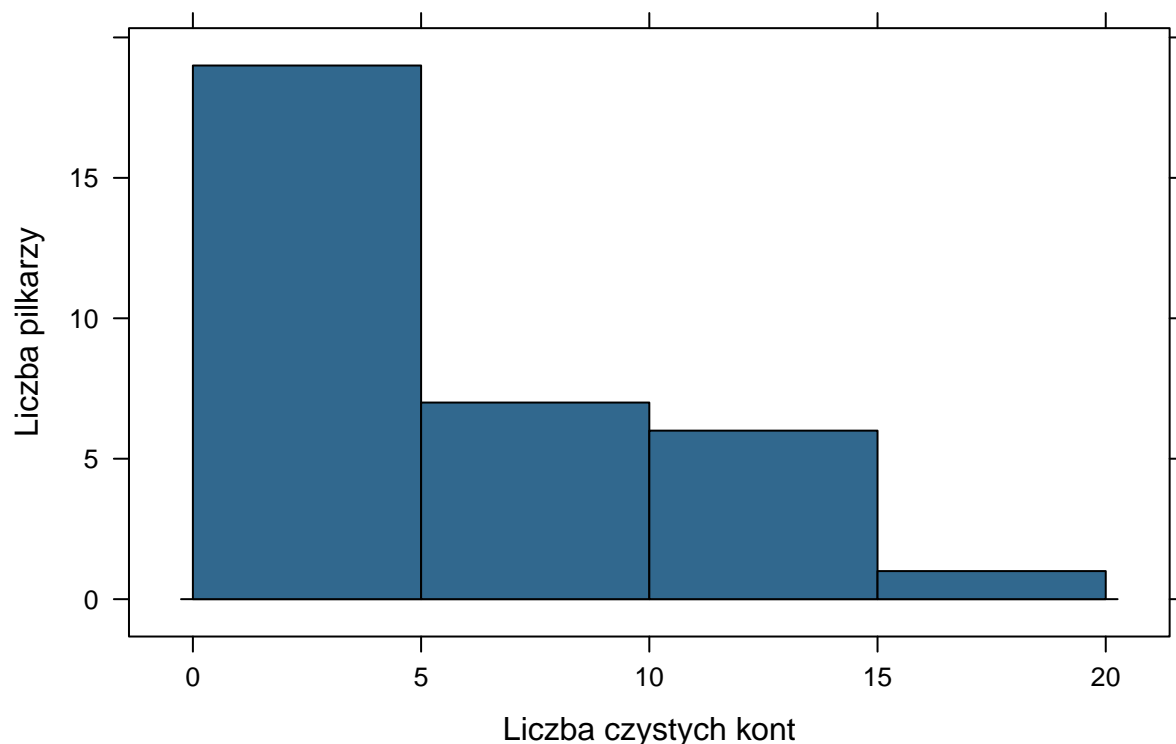
Bramkarze

Powszechnie wiadomo, że bramkarze są odpowiedzialni za bronienie dostępu do własnej bramki. Z tego względu, nie ma wielu statystyk dotyczącego tego aspektu gry. Poniższy rozdział będzie zawierał analizę takich elementów jak, ilość czystych kont, liczba straconych bramek, błędy prowadzące do utraty bramki, czy liczba obronionych strzałów wśród bramkarzy. Zostanie również poruszona kwestia gry nogami bramkarzy oraz ich skuteczności. Jest to aspekt, na który coraz częściej zwracają uwagę trenerzy w obecnych czasach.

Analiza skuteczności bronienia własnej bramki

W tym podrozdziale zostanie przedstawiona analiza skuteczności bramkarzy w bronieniu dostępu do własnej bramki. Początkowo, został wygenerowany wykres przedstawiający histogram czystych kont zawodników (sytuacja gry bramkarz nie straci ani jednej bramki w meczu), czyli najważniejszej statystyki defensywnej jeśli chodzi o bramkarzy.

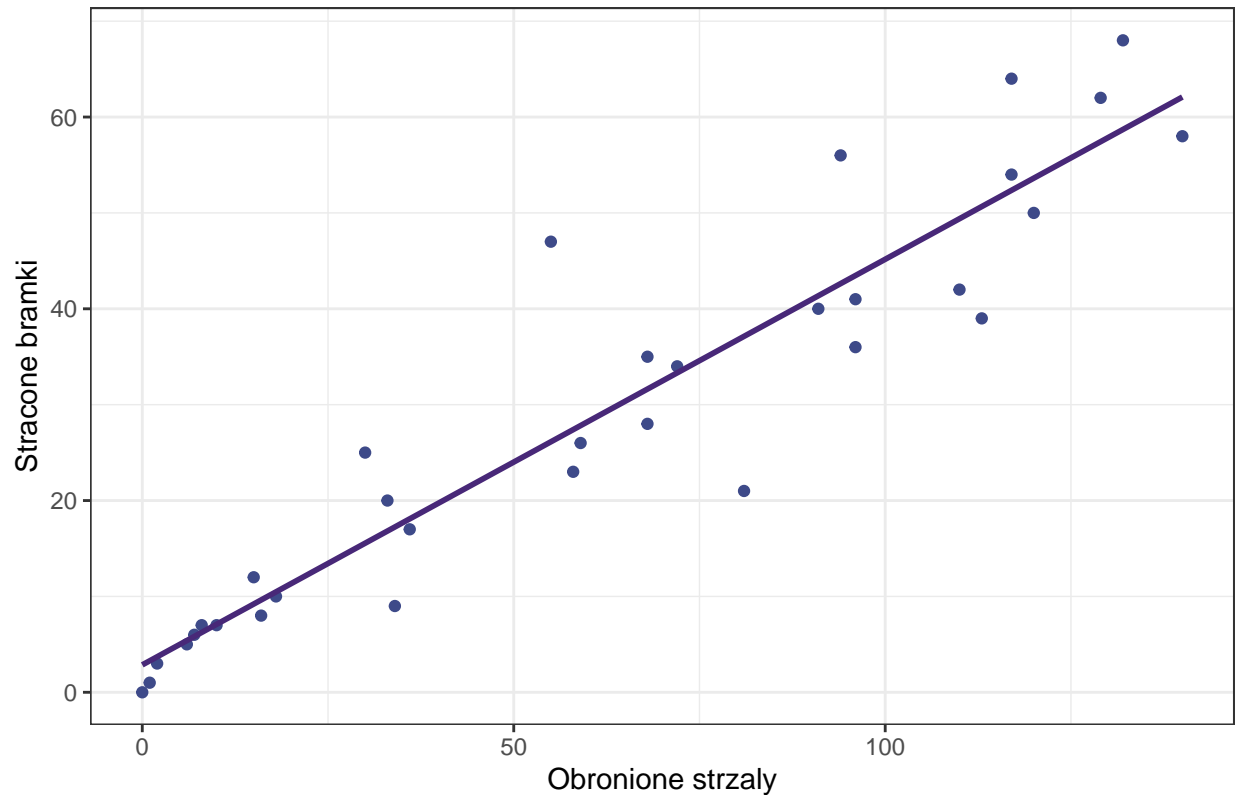
Histogram czystych kont bramkarzy



Na powyższym wykresie można zauważyć, że najliczniejsza liczba bramkarzy zachowała czyste konto maksymalnie pięć razy w ciągu sezonu. Może to być skutkiem tego, że jest nieznaczna liczba bramkarzy, którzy stosunkowo rzadko występowali na boisku. Widzimy również, że dosłownie kilku bramkarzy zachowało czyste konto od 10 do 20 razy. Prawdopodobnie, są to zawodnicy najlepszych drużyn w lidze, które stosunkowo rzadko przegrywają, co wiąże się z małą liczbą traconych przez nich bramek. Kolejno, zostanie przeanalizowana zależność pomiędzy liczbą straconych bramek, a obronionymi strzałami.

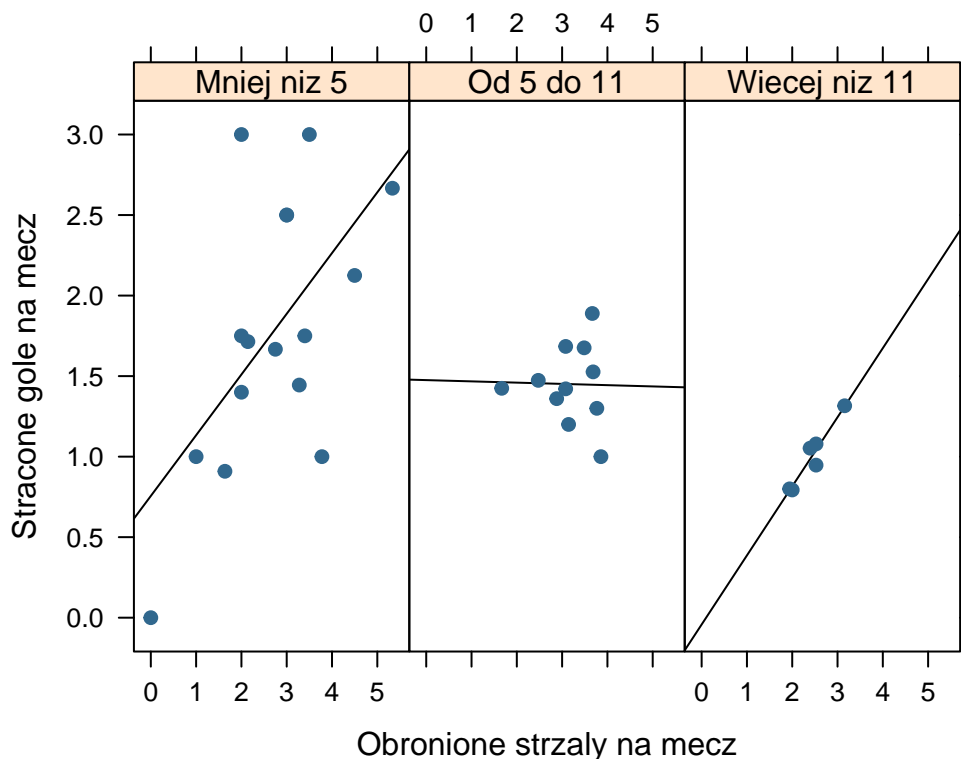
```
## 'geom_smooth()' using formula 'y ~ x'
```

Stracone bramki a obronione strzaly



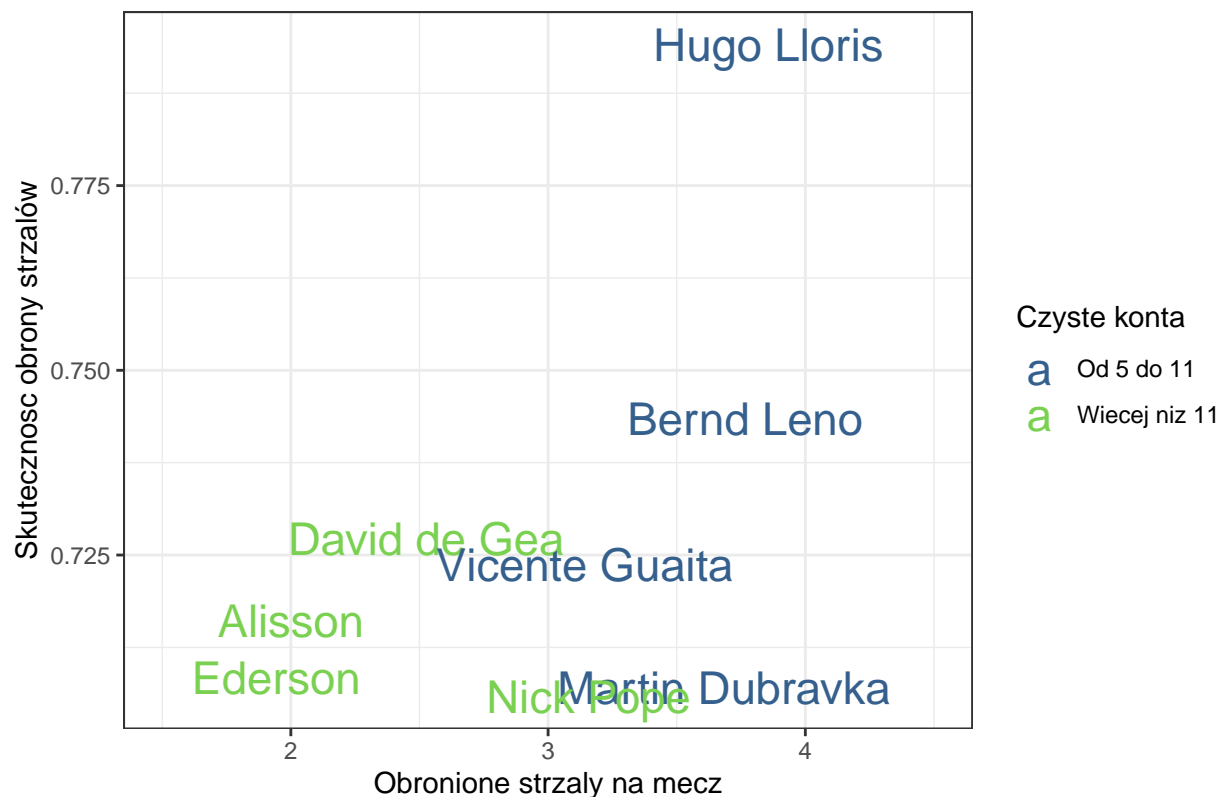
Powyższy wykres dostarcza ciekawych wniosków. Naturalnie moglibyśmy pomyśleć, że im większa liczba strzałów, tym mniej straconych bramek przez bramkarza. Nie jest to jednak prawdą, ponieważ może to oznaczać że na bramkę zawodnika pada zdecydowanie więcej strzałów, z którymi musi on sobie radzić, co powoduje większe prawdopodobieństwo utraty bramki. Inną ważną kwestią do zanotowania jest fakt, że bramkarze którzy obronili małą liczbę strzałów, prawdopodobnie wystąpili w mniejszej ilości spotkań. To powoduje, że liczba utraconych przez nich bramek może być niższa. Z tego powodu została przeprowadzona analiza liczby straconych bramek na mecz oraz obronionych strzałów na mecz w podziale na liczbę czystych kont. W tym celu zostanie utworzona nowa zmienna kategoryzująca bazująca na statystyce czystych kont.

Ilość straconych goli od obronionych strzałów w podziale ilość czystych



Powyższe wykresy dostarczają bardzo ciekawych wniosków. Mianowicie, bramkarze którzy zachowali czyste konto więcej niż 11 razy wcale nie musieli radzić sobie ze znacznie mniejszą ilością strzałów na mecz. W ich kategorii, widać pozytywną korelację pomiędzy liczbą strzałów na mecz a liczbą straconych goli na mecz. Podobna sytuacja występuje u bramkarzy, którzy zachowali czyste konto mniej niż 5 razy. W tej kategorii, zawodnicy cieszyli się największą niechlubną statystyką liczby straconych bramek na mecz. Dwóch bramkarzy traciło średnio aż 3 bramki na mecz. W kategorii pośredniej, gdzie bramkarzy zachowywali od 5 do 11 czystych kont w ciągu sezonu nie widać korelacji pomiędzy liczbą obronionych strzałów na mecz, a straconymi bramkami na mecz. W ostatniej części tego podrozdziału zostanie wygenerowany wykres, przedstawiający bramkarzy z największym procentem obronionych strzałów. Do analizy zostali włączeni jedynie bramkarze, którzy wystąpili w przynajmniej 10 meczach.

Bramkarze z największym procentem obronionych strzałów

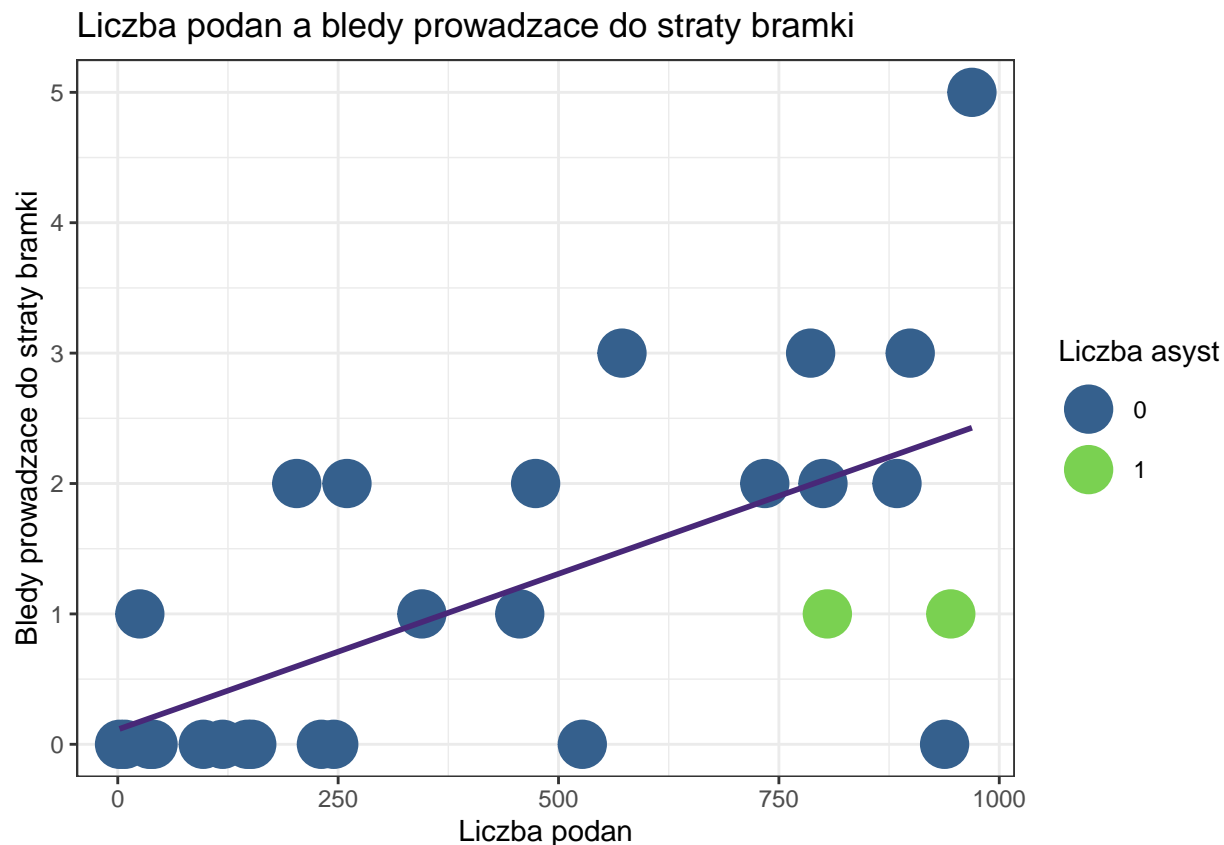


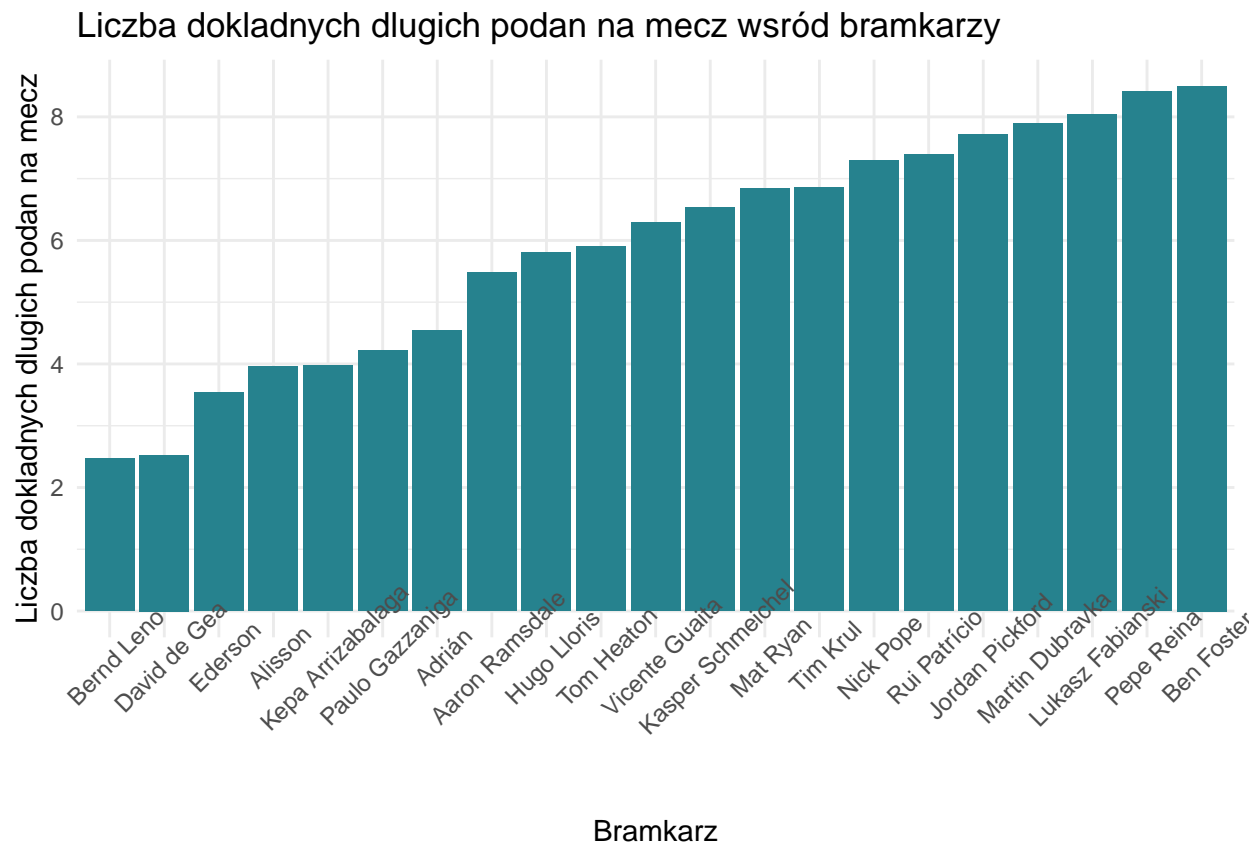
Jak widać na powyższym wykresie, zawodnikiem który znacznie wyróżniał się od swoich konkurentów w tym aspekcie gry był Hugo Lloris. Mógł się on pochwalić największą skutecznością obrony strzałów, przy największej ilości oddanych strzałów na jego bramkę wśród analizowanych zawodników. Co ciekawe, zachował on jedynie od 5 do 11 czystych kont w ciągu sezonu. Więcej czystych kont zachowali tacy bramkarze jak David de Gea, Alisson, Ederson czy Nick Pope.

Gra nogami bramkarzy

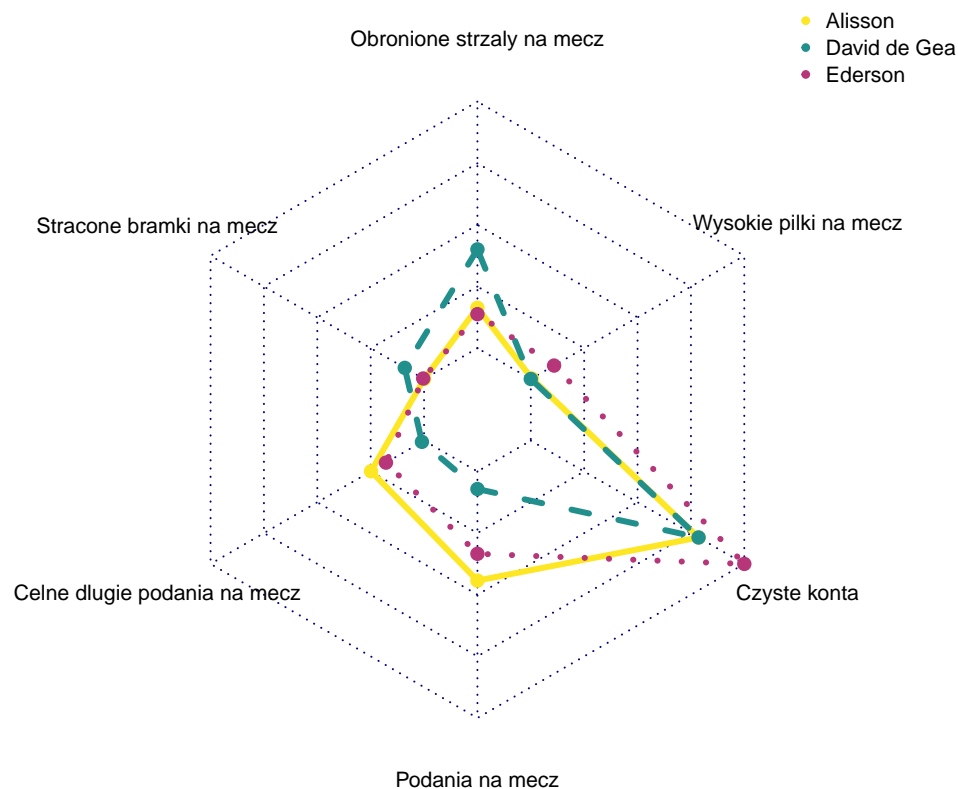
Tak jak już zostało wspomniane na wstępie do tego rozdziału, element gry nogami wśród bramkarzy staje się coraz istotniejszy we współczesnym futbolu. Poniższy podrozdział ma na celu przybliżenie tego aspektu gry wśród bramkarzy ligi angielskiej. W pierwszej kolejności został przedstawiony wykres, ukazujący zależność pomiędzy liczbą podań bramkarza a ilością błędów prowadzących do straty bramki oraz asyst.

```
## 'geom_smooth()' using formula 'y ~ x'
```

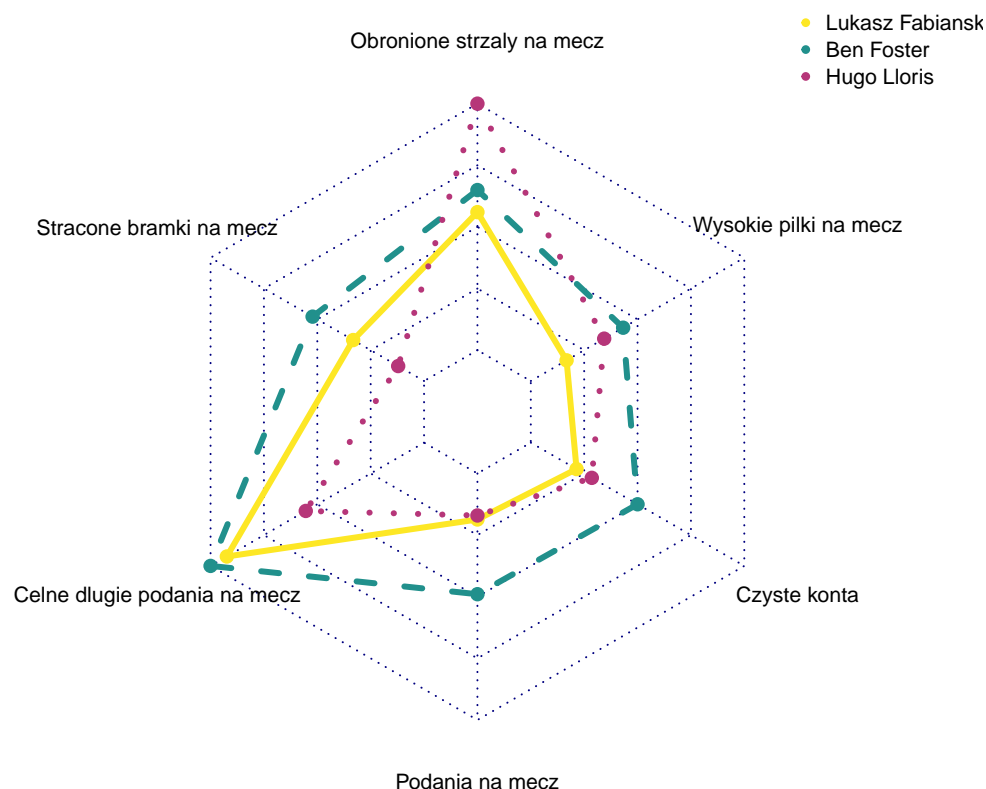





Widzimy na powyższym wykresie że trzech bramkarzy ma ponad 8 celnych długich podań na mecz. Możemy się spodziewać, że są to bramkarze drużyn, które lubią grać z kontrataku i szybko przenosić piłkę na połowę rywali. Czterech bramkarzy z najniższą liczbą długich podań na mecz, fani piłki nożnej mogą kojarzyć z drużyn grających krótkimi podaniami i wyprowadzającymi piłkę spod własnej bramki. Na koniec analizy bramkarzy zostaną przedstawione cztery wykresy radarowe dla wybranych bramkarzy.



Wykresy radarowe są bardzo użyteczną formą wizualizacji statystyki zawodników i porównywanie ich. Powyższe wykresy dostarczają bardzo wartościowych informacji na temat konkretnych zawodników i ich skuteczności w grze. W pierwszej kolejności wszystkie powyżej przedstawione statystyki zostały znormalizowane przy wykorzystaniu następującej funkcji: $y(x) = (x - \min(\mathbf{X})) / (\max(\mathbf{X}) - \min(\mathbf{X}))$, gdzie x należy do wektora \mathbf{X} . Na podstawie pierwszego wykresu, możemy zauważyć że Ederson był najlepszym bramkarzem pod względem liczby czystych kont w porównaniu do dwóch innych bramkarzy, Alissona oraz Davida de Gei. Natomiast jeśli chodzi o liczbę podań na mecz, najlepszy w tej kategorii w tej trójce był Alisson. David de Gea przewodzi w kategorii obronionych strzałów na mecz.



Jeśli chodzi o drugi wykres, widać zdecydowaną przewagę Hugo Llorisa w kategorii liczby obronionych strzałów na mecz nad dwójką innych bramkarzy, Łukasza Fabiańskiego oraz Bena Fostera. Przewodzi on również pod względem liczby straconych bramek na mecz, ponieważ w tej statystyce im mniejsza liczba tym pozytywniej świadczy to o zawodniku.

Analiza statystyk obrońców

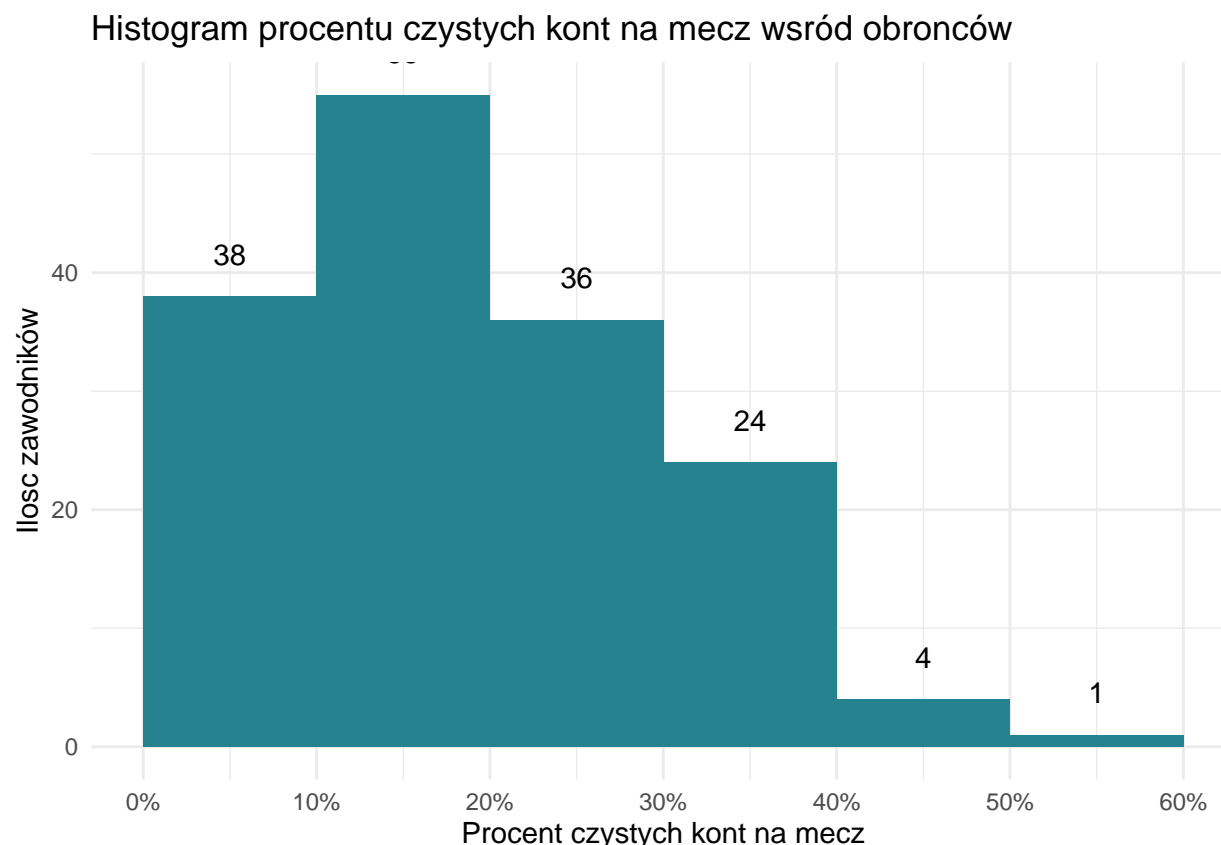
W tym rozdziale zostanie przeprowadzona analiza statystyk piłkarzy, występujących na pozycji obrońcy. Najważniejszą rolą obrońcy jest uniemożliwienie stworzenia dogodnej sytuacji do zdobycia bramki przez przeciwników. Dlatego, ten rozdział w głównej mierze będzie poświęcony analizie statystyk, mówiących o skuteczności bronięcia dostępu do własnej bramki, czy agresywności. W obecnych czasach, w których coraz częściej jest promowany ofensywny futbol, nawet obrońcy są odpowiedzialni za tworzenie okazji do zdobycia bramki. Z tego powodu, zostanie przeprowadzona również krótka analiza skuteczności obrońców pod bramką rywala.

Analiza defensywnych statystyk obrońców

Poniższy podrozdział będzie zawierał analizę defensywnych statystyk obrońców, takich jak liczba czystych kont, liczba pojedynków oraz ich skuteczność, czy liczba przechwyty. W bazie znajduje się 161 obrońców. Poniżej zaprezentowano histogram procentu liczby czystych kont w sezonie wśród obrońców.

```
## Warning: Removed 3 rows containing non-finite values (stat_bin).
```

```
## Warning: Removed 3 rows containing non-finite values (stat_bin).
```



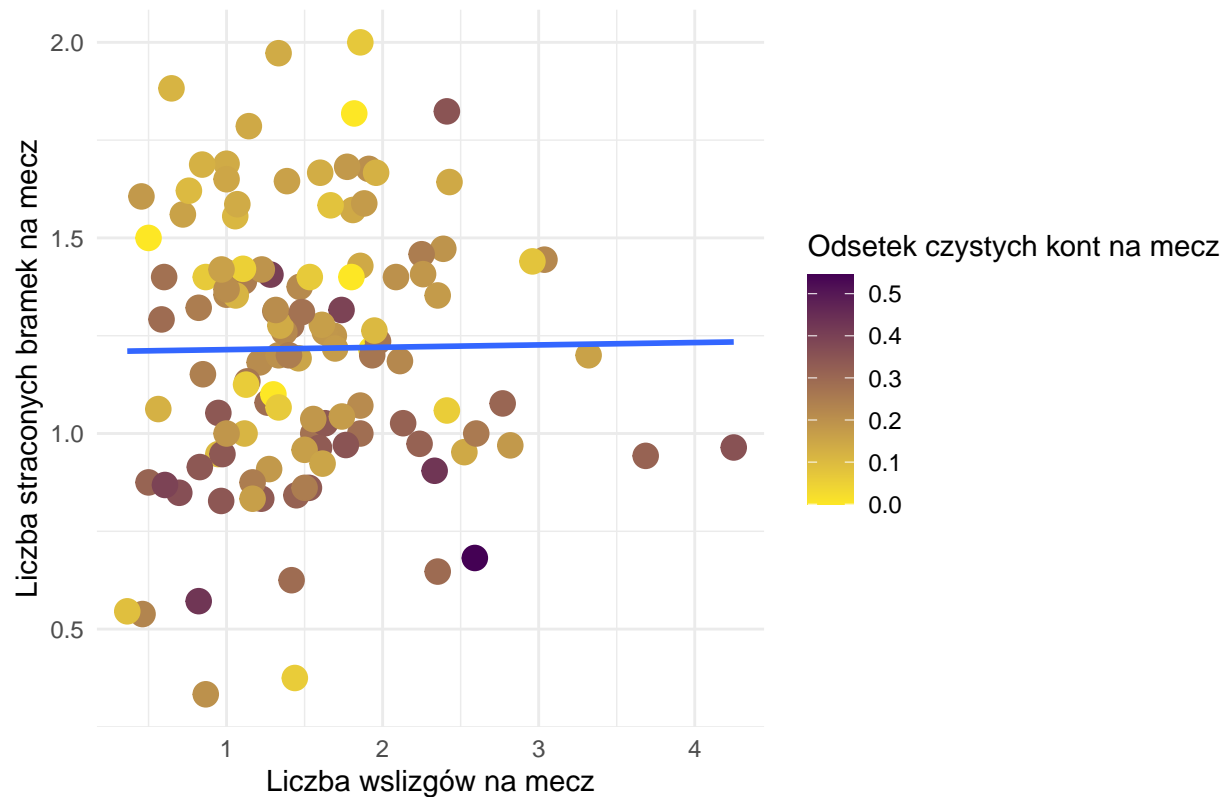
Powyższy wykres dostarcza informacji, że wartość odsetka czystych kont na mecz w sezonie najczęściej wynosiła z przedziału 10% - 20%. Aż 55 zawodników zaliczało się do tego przedziału. Tylko jeden zawodnik co drugi mecz ukończył z czystym kontem. Następnie została zwizualizowana zależność pomiędzy liczbą pojedynków na mecz a liczbą straconych bramek na mecz w podziale na odsetek czystych kont na mecz. Poniższa analiza uwzględnia obrońców, którzy rozegrali przynajmniej 10 spotkań w ciągu sezonu.

```
## 'geom_smooth()' using formula 'y ~ x'
```

```
## Warning: Removed 3 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 3 rows containing missing values (geom_point).
```

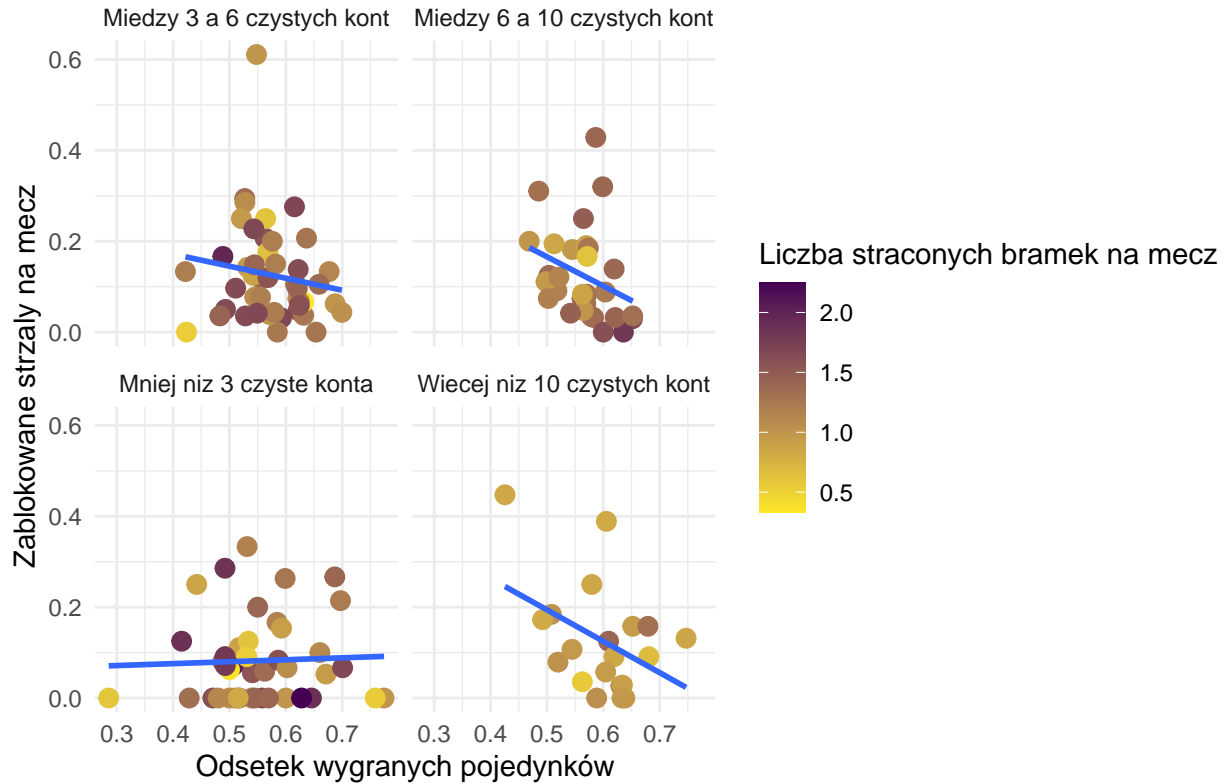
Liczba wślizgów na mecz a liczba straconych bramek na mecz w podziale r



Na podstawie powyższego wykresu, można wysnuć wniosek że nie ma zależności pomiędzy liczbą wślizgów na mecz a liczbą straconych bramek na mecz. Widać natomiast pewną ujemną zależność pomiędzy liczbą straconych bramek na mecz a odsetkiem czystych kont na mecz, co wydaje się być oczywistą zależnością. Niżej osadzone punkty mają raczej ciemniejszą barwę niż te powyżej, co sugeruje że wraz z mniejszą liczbą straconych bramek na mecz odsetek czystych kont na mecz wzrasta. Następnie przeanalizowano wpływ odsetka wygranych pojedynków na liczbę zablokowanych strzałów na mecz.

```
## 'geom_smooth()' using formula 'y ~ x'
```

Zablokowane strzale na mecz a skuteczność pojedynków oraz liczbe stracone



Powyższy wykres dostarcza wiele różnych wniosków na temat gry obronnej zawodników. W pierwszej kolejności, w trzech z czterech kategorii widać ujemną zależność pomiędzy skutecznością pojedynków a liczbą zablokowanych strzałów na mecz. Natomiast obrońcy, którzy zachowali czyste konta mniej niż 3 razy wyłamują się z tego schematu. W ich przypadku tej zależności nie widać. Nie widać zależności pomiędzy liczbą straconych bramek na mecz a liczbą zablokowanych strzałów na mecz. Tak samo jak pomiędzy liczbą straconych bramek a skutecznością pojedynków.