

League of Legends Diamond Ranked Games (10 min)

Etap pierwszy: Zrozumienie problemu + Zrozumienie danych

Ogólny opis zbioru danych.....	1
Potencjalne zastosowania:.....	1
Cele eksploracji danych.....	2
Cel eksploracji.....	2
Kryteria sukcesu.....	2
Charakterystyka zbioru danych.....	2
Opis atrybutów.....	2
Eksploracyjna analiza danych.....	4
Rozkłady wartości atrybutów.....	4
Rozkład zmiennej docelowej.....	15
Korelacje pomiędzy atrybutami.....	16
Macierz korelacji.....	19
Ocena jakości danych.....	20
Brakujące dane.....	20
Niespójne dane.....	20
Dane niezrozumiałe.....	20
Wnioski.....	30
Ocena realizacji celów.....	30
Rewizje celów.....	30
Dodatkowe cele.....	30
Podsumowanie.....	31

Ogólny opis zbioru danych

Zbiór danych zawiera informacje o grach z wysokiej dywizji diamentu w League of Legends, koncentrując się na statystykach z pierwszych 10 minut rozgrywki. Dane te są szczególnie wartościowe, ponieważ pokazują wczesne etapy gry, które często determinują jej ostateczny wynik.

Źródło: [Kegggle](#)

Kategoria: Gry, e-sport

Zawartość: Statystyki z gier rankingowych Diamond w League of Legends

Horyzont czasowy: Pierwsze 10 minut każdej gry

Zmienna docelowa: Wynik gry (wygrana/przegrana)

Potencjalne zastosowania:

- Analiza czynników we wczesnej fazie gry wpływających na zwycięstwo
- Badanie wpływu pierwszych 10 minut gry na ostateczny wynik
- Tworzenie modeli predykcyjnych do przewidywania wyniku gry na podstawie początkowych statystyk
- Badanie wzorców strategicznych w grach wysokiej rangi

Cele eksploracji danych

Cel eksploracji

- Identyfikacja kluczowych czynników z pierwszych 10 minut gry, które najsilniej wpływają na końcowy wynik rozgrywki.
- Budowa skutecznego modelu predykcyjnego, który na podstawie statystyk z pierwszych 10 minut pozwoli przewidzieć zwycięzcę meczu.
- Odkrycie istotnych wzorców i zależności strategicznych w początkowej fazie gry na wysokim poziomie rozgrywek.

Kryteria sukcesu

- Wyodrębnienie statystycznie istotnych predyktorów wyniku gry
- Stworzenie modelu o dokładności predykcji powyżej 70%
- Wykrycie nietrywialnych zależności między zmiennymi

Charakterystyka zbioru danych

Parametr	Wartość
Format	CSV (Comma-Separated Values)
Liczba rekordów	9879 gier
Liczba atrybutów	40 zmiennych (włączając ID gry i zmienną docelową)
Struktura	Jednolity zbiór danych, gdzie każdy wiersz reprezentuje jedną rozgrywkę
Zmienna docelowa	blueWins (czy niebieska drużyna wygrała grę - wartości 0/1)

Opis atrybutów

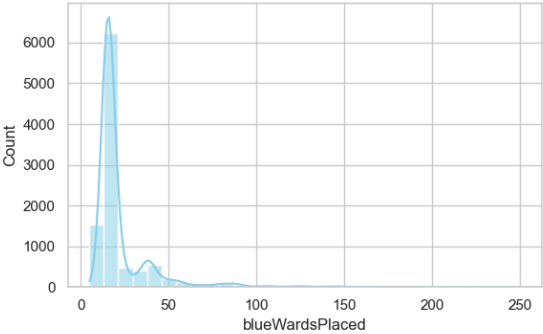
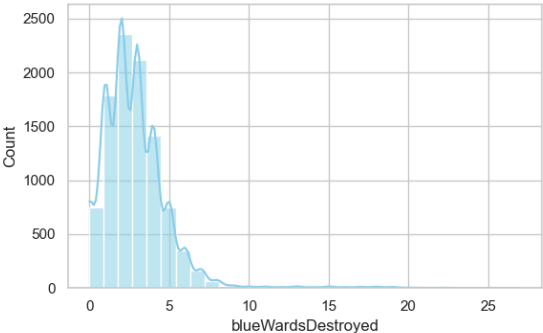
Grupa atrybutów	Nazwa atrybutu	Typ	Opis	Jednostka/Zakres
Identyfikacja	gameId	Liczbowy	Unikalny identyfikator gry	Dodatnia liczba całkowita
Wynik	blueWins	Nominalny	Czy drużyna niebieska wygrała mecz	0=nie, 1=tak
Zabójstwa i śmierci	blueKills / redKills	Liczbowy	Liczba zabójstw zdobytych przez drużynę	Nieujemna liczba całkowita
	blueDeaths / redDeaths	Liczbowy	Liczba śmierci w drużynie	Nieujemna liczba całkowita
	blueAssists / redAssists	Liczbowy	Liczba asyst zdobytych przez drużynę	Nieujemna liczba całkowita
	blueFirstBlood / redFirstBlood	Nominalny	Czy drużyna zdobyła pierwsze zabójstwo	0=nie, 1=tak
Ekonomia	blueTotalGold / redTotalGold	Liczbowy	Całkowita ilość złota zdobytego przez drużynę	Nieujemna liczba całkowita
	blueGoldPerMin / redGoldPerMin	Liczbowy	Średnia ilość złota zdobywana na minutę	Liczba rzeczywista
	blueGoldDiff	Liczbowy	Różnica w złocie między drużyną niebieską a czerwoną	Liczba całkowita

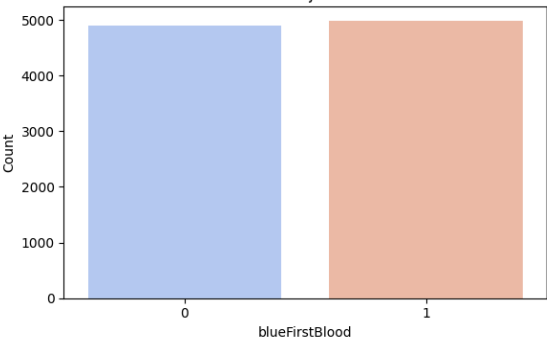
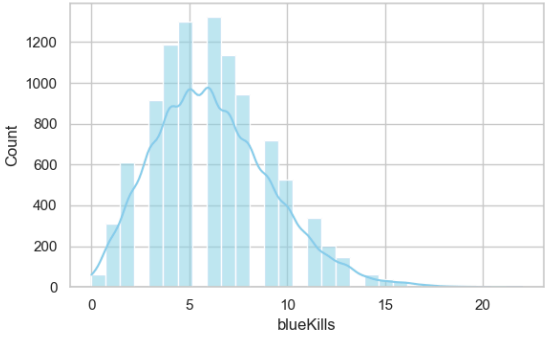
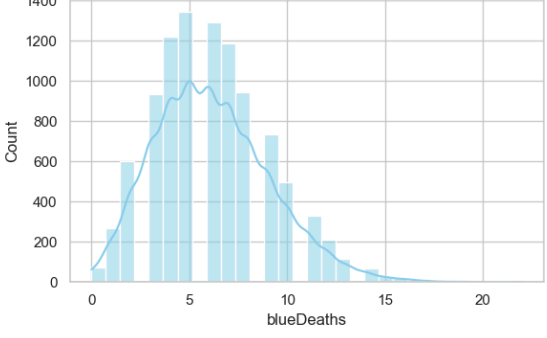
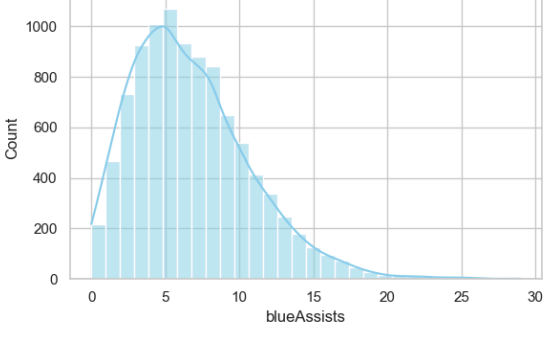
Rozwój postaci	blueAvgLevel / redAvgLevel	Liczbowy	Średni poziom bohaterów w drużynie	Liczba rzeczywista > 1
	blueTotalExperience / redTotalExperience	Liczbowy	Całkowita ilość doświadczenia zdobytego przez drużynę	Nieujemna liczba całkowita
	blueExperienceDiff	Liczbowy	Różnica w doświadczeniu między drużyną niebieską a czerwoną	Liczba całkowita
Farmienie	blueTotalMinionsKilled / redTotalMinionsKilled	Liczbowy	Całkowita liczba zabitych minionów	Nieujemna liczba całkowita
	blueTotalJungleMinionsKilled / redTotalJungleMinionsKilled	Liczbowy	Całkowita liczba zabitych potworów w dżungli	Nieujemna liczba całkowita
	blueCSPerMin / redCSPerMin	Liczbowy	Średnia liczba minionów zabijanych na minutę	Nieujemna liczba rzeczywista
Obiekty	blueTowersDestroyed / redTowersDestroyed	Liczbowy	Liczba zniszczonych wież przeciwnika	Liczba całkowita z przedziału 0-3
	blueEliteMonsters / redEliteMonsters	Liczbowy	Liczba zabitych elitarnych potworów	Nieujemna liczba całkowita
	blueDragons / redDragons	Liczbowy	Liczba zabitych smoków	Liczba całkowita z przedziału 0-2
	blueHeralds / redHeralds	Liczbowy	Liczba zabitych heroldów	Liczba całkowita z przedziału 0-1
Wizja	blueWardsPlaced / redWardsPlaced	Liczbowy	Liczba postawionych totemów (wardów)	Nieujemna liczba całkowita
	blueWardsDestroyed / redWardsDestroyed	Liczbowy	Liczba zniszczonych totemów	Nieujemna liczba całkowita

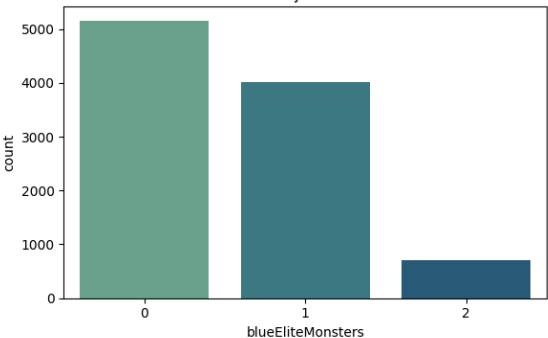
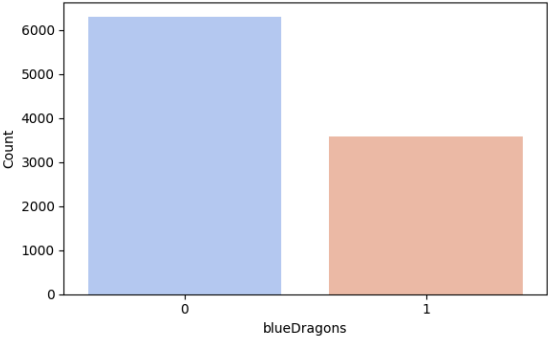
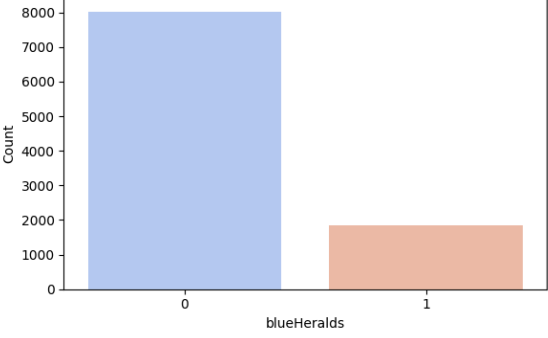
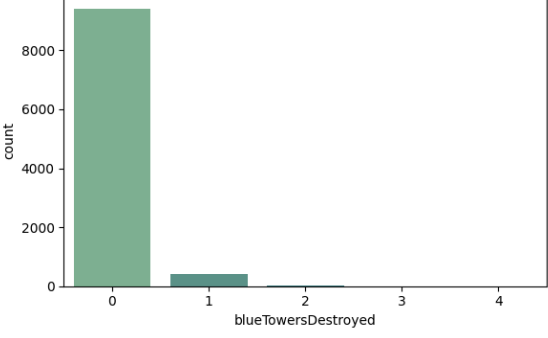
	ed		przeciwnika	
--	----	--	-------------	--

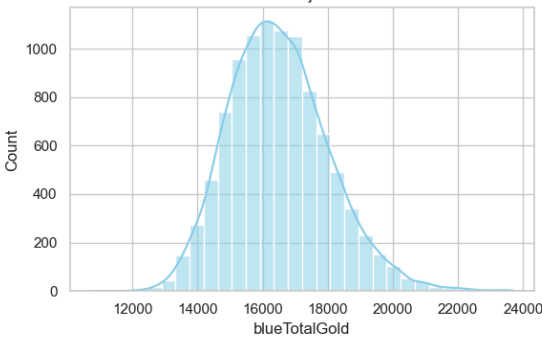
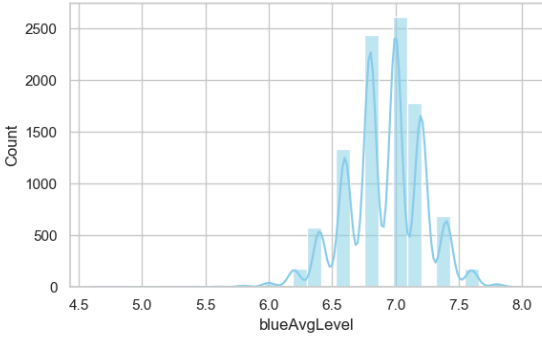
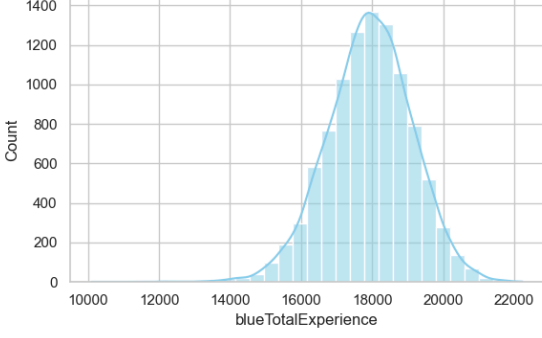
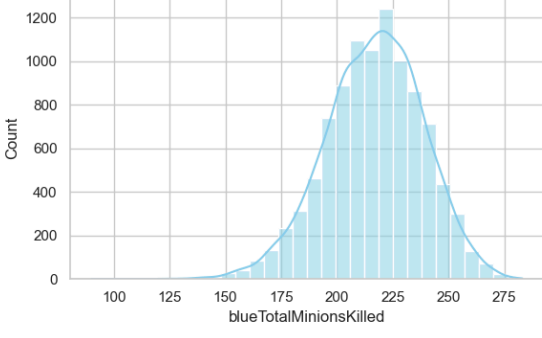
Eksploracyjna analiza danych

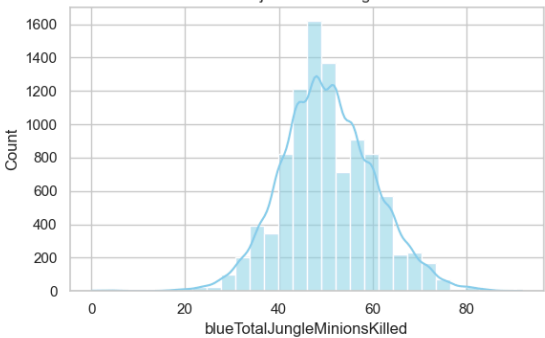
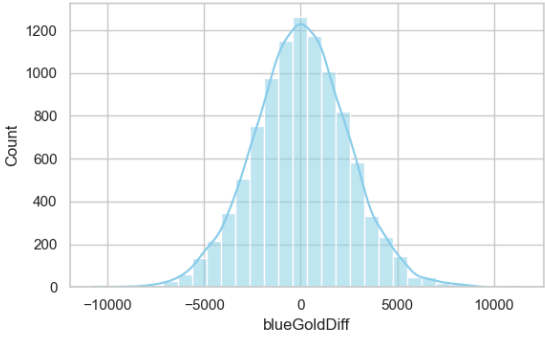
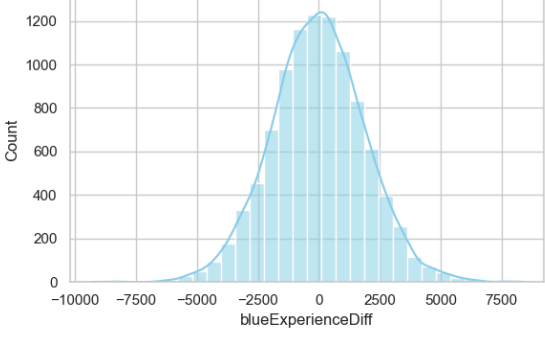
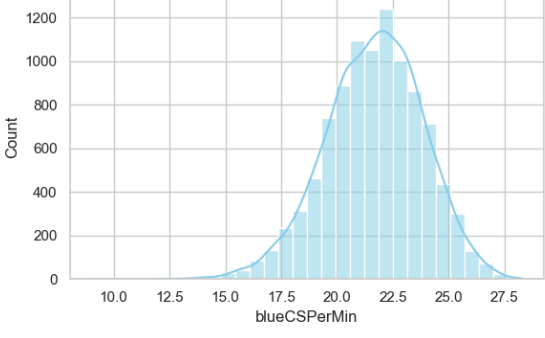
Rozkłady wartości atrybutów

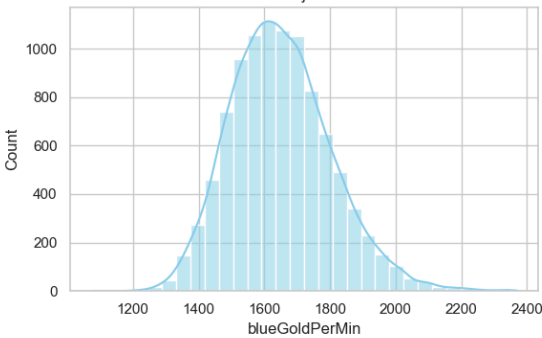
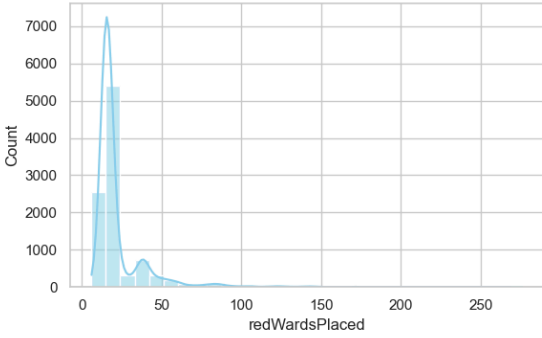
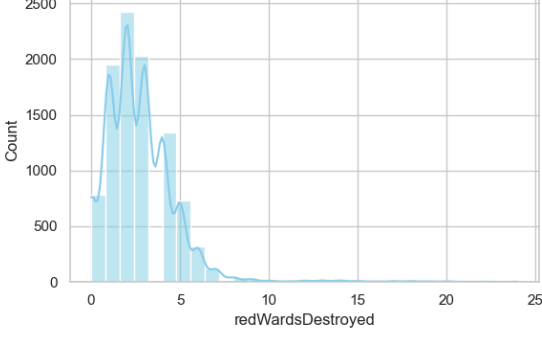
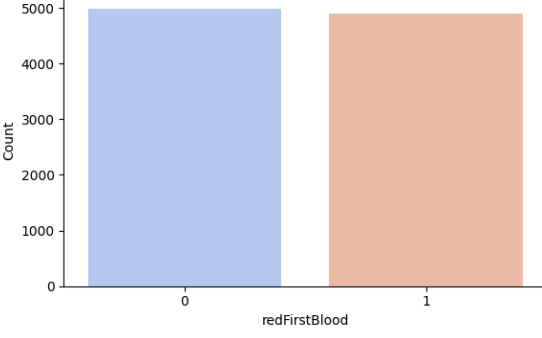
Atrybut	Rozkłady wartości atrybutów
blueWardsPlaced	<div><p>Rozkład zmiennej: blueWardsPlaced</p></div>
blueWardsDestroyed	<div><p>Rozkład zmiennej: blueWardsDestroyed</p></div>

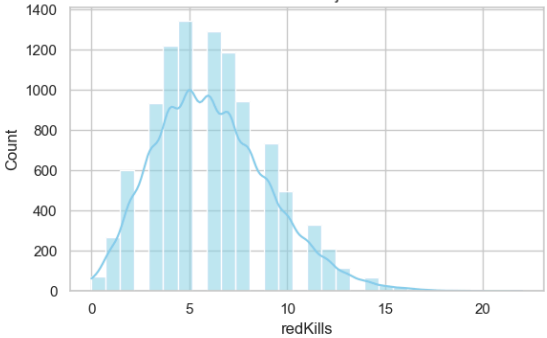
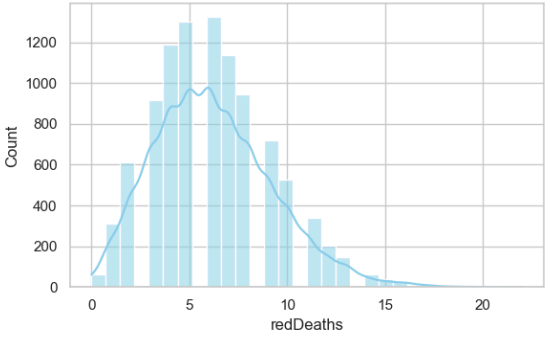
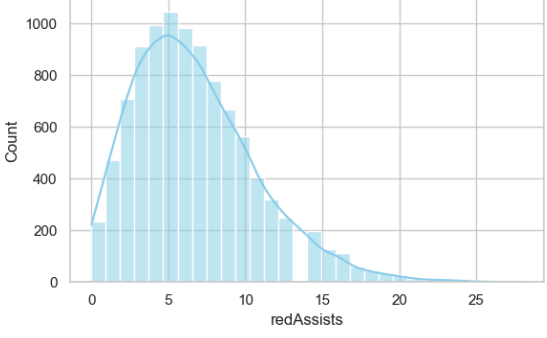
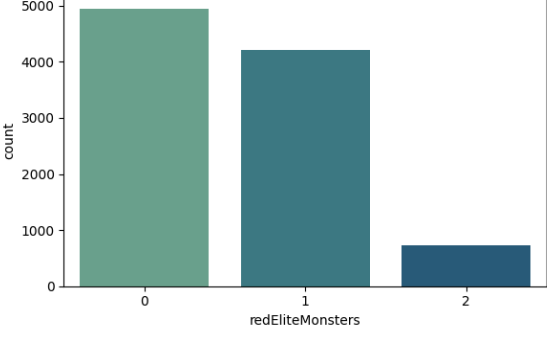
blueFirstBlood	<p>Rozkład zmiennej: blueFirstBlood</p> 
blueKills	<p>Rozkład zmiennej: blueKills</p> 
blueDeaths	<p>Rozkład zmiennej: blueDeaths</p> 
blueAssists	<p>Rozkład zmiennej: blueAssists</p> 

blueEliteMonsters	<p>Rozkład zmiennej: blueEliteMonsters</p>  <p>A histogram showing the distribution of the variable 'blueEliteMonsters'. The x-axis is labeled 'blueEliteMonsters' and has tick marks at 0, 1, and 2. The y-axis is labeled 'count' and ranges from 0 to 5000. There are three bars: a green bar at 0 with a count of approximately 5100, a teal bar at 1 with a count of approximately 4000, and a dark blue bar at 2 with a count of approximately 700.</p> <table border="1"><thead><tr><th>blueEliteMonsters</th><th>count</th></tr></thead><tbody><tr><td>0</td><td>5100</td></tr><tr><td>1</td><td>4000</td></tr><tr><td>2</td><td>700</td></tr></tbody></table>	blueEliteMonsters	count	0	5100	1	4000	2	700				
blueEliteMonsters	count												
0	5100												
1	4000												
2	700												
blueDragons	<p>Rozkład zmiennej: blueDragons</p>  <p>A histogram showing the distribution of the variable 'blueDragons'. The x-axis is labeled 'blueDragons' and has tick marks at 0 and 1. The y-axis is labeled 'Count' and ranges from 0 to 6000. There are two bars: a light blue bar at 0 with a count of approximately 6200, and an orange bar at 1 with a count of approximately 3600.</p> <table border="1"><thead><tr><th>blueDragons</th><th>Count</th></tr></thead><tbody><tr><td>0</td><td>6200</td></tr><tr><td>1</td><td>3600</td></tr></tbody></table>	blueDragons	Count	0	6200	1	3600						
blueDragons	Count												
0	6200												
1	3600												
blueHeralds	<p>Rozkład zmiennej: blueHeralds</p>  <p>A histogram showing the distribution of the variable 'blueHeralds'. The x-axis is labeled 'blueHeralds' and has tick marks at 0 and 1. The y-axis is labeled 'Count' and ranges from 0 to 8000. There are two bars: a light blue bar at 0 with a count of approximately 8000, and an orange bar at 1 with a count of approximately 1900.</p> <table border="1"><thead><tr><th>blueHeralds</th><th>Count</th></tr></thead><tbody><tr><td>0</td><td>8000</td></tr><tr><td>1</td><td>1900</td></tr></tbody></table>	blueHeralds	Count	0	8000	1	1900						
blueHeralds	Count												
0	8000												
1	1900												
blueTowersDestroyed	<p>Rozkład zmiennej: blueTowersDestroyed</p>  <p>A histogram showing the distribution of the variable 'blueTowersDestroyed'. The x-axis is labeled 'blueTowersDestroyed' and has tick marks at 0, 1, 2, 3, and 4. The y-axis is labeled 'count' and ranges from 0 to 8000. There are five bars: a green bar at 0 with a count of approximately 9000, a teal bar at 1 with a count of approximately 500, a dark blue bar at 2 with a count of approximately 100, and very small bars at 3 and 4.</p> <table border="1"><thead><tr><th>blueTowersDestroyed</th><th>count</th></tr></thead><tbody><tr><td>0</td><td>9000</td></tr><tr><td>1</td><td>500</td></tr><tr><td>2</td><td>100</td></tr><tr><td>3</td><td>50</td></tr><tr><td>4</td><td>50</td></tr></tbody></table>	blueTowersDestroyed	count	0	9000	1	500	2	100	3	50	4	50
blueTowersDestroyed	count												
0	9000												
1	500												
2	100												
3	50												
4	50												

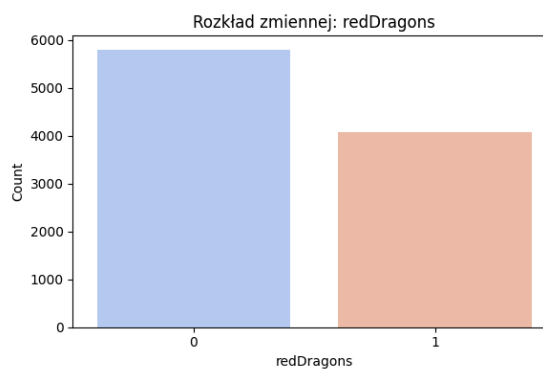
blueTotalGold	<p>Rozkład zmiennej: blueTotalGold</p> 
blueAvgLevel	<p>Rozkład zmiennej: blueAvgLevel</p> 
blueTotalExperience	<p>Rozkład zmiennej: blueTotalExperience</p> 
blueTotalMinionsKilled	<p>Rozkład zmiennej: blueTotalMinionsKilled</p> 

blueTotalJungleMinionsKilled	<p>Rozkład zmiennej: blueTotalJungleMinionsKilled</p> 
blueGoldDiff	<p>Rozkład zmiennej: blueGoldDiff</p> 
blueExperienceDiff	<p>Rozkład zmiennej: blueExperienceDiff</p> 
blueCSPerMin	<p>Rozkład zmiennej: blueCSPerMin</p> 

blueGoldPerMin	<p>Rozkład zmiennej: blueGoldPerMin</p>  <p>A histogram showing the distribution of blueGoldPerMin. The x-axis is labeled 'blueGoldPerMin' and ranges from 1200 to 2400. The y-axis is labeled 'Count' and ranges from 0 to 1000. The distribution is roughly bell-shaped, centered around 1600. A blue normal distribution curve is overlaid on the histogram bars.</p>
redWardsPlaced	<p>Rozkład zmiennej: redWardsPlaced</p>  <p>A histogram showing the distribution of redWardsPlaced. The x-axis is labeled 'redWardsPlaced' and ranges from 0 to 250. The y-axis is labeled 'Count' and ranges from 0 to 7000. The distribution is highly right-skewed, with a very high peak near 0 and a long tail extending towards 250. A blue normal distribution curve is overlaid, showing it does not fit the data well.</p>
redWardsDestroyed	<p>Rozkład zmiennej: redWardsDestroyed</p>  <p>A histogram showing the distribution of redWardsDestroyed. The x-axis is labeled 'redWardsDestroyed' and ranges from 0 to 25. The y-axis is labeled 'Count' and ranges from 0 to 2500. The distribution is right-skewed, with a peak around 2 and a tail extending to 25. A blue normal distribution curve is overlaid.</p>
redFirstBlood	<p>Rozkład zmiennej: redFirstBlood</p>  <p>A bar chart showing the distribution of redFirstBlood. The x-axis is labeled 'redFirstBlood' and has two categories: 0 and 1. The y-axis is labeled 'Count' and ranges from 0 to 5000. The bar for category 0 is blue and has a count of approximately 5000. The bar for category 1 is orange and has a count of approximately 4800.</p>

redKills	<p>Rozkład zmiennej: redKills</p> 
redDeaths	<p>Rozkład zmiennej: redDeaths</p> 
redAssists	<p>Rozkład zmiennej: redAssists</p> 
redEliteMonsters	<p>Rozkład zmiennej: redEliteMonsters</p> 

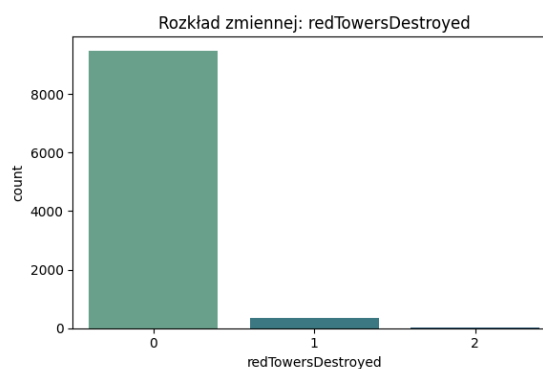
redDragons



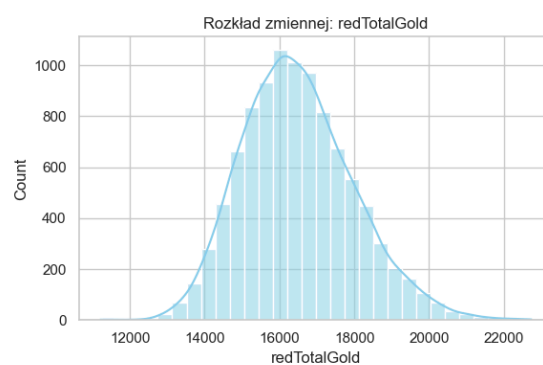
redHeralds

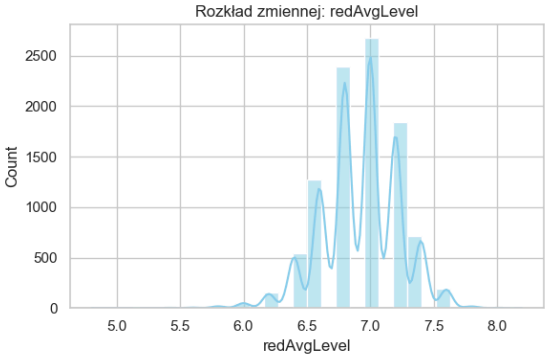
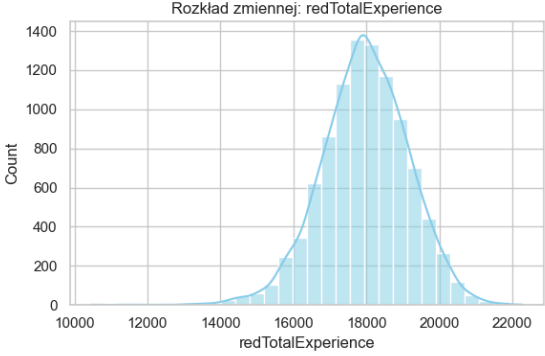
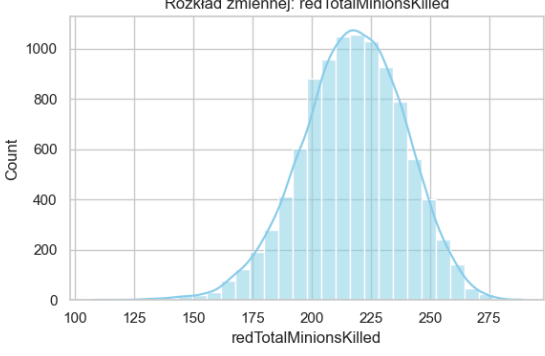
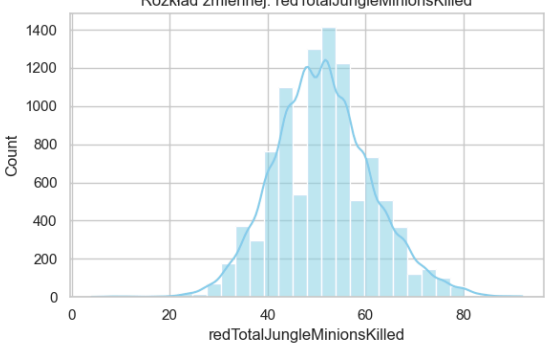


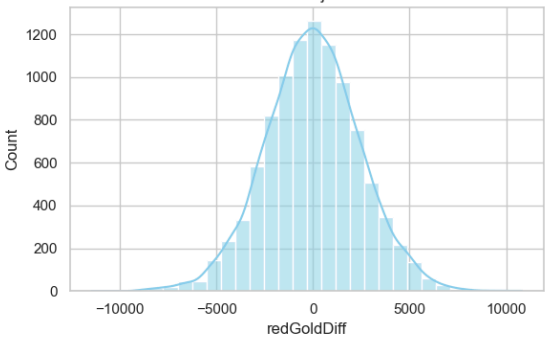
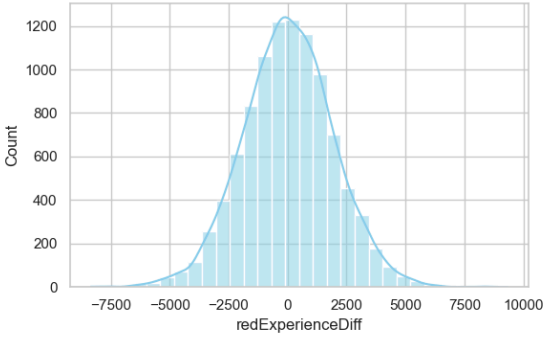
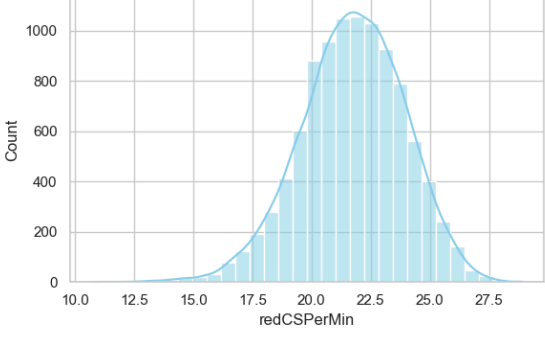
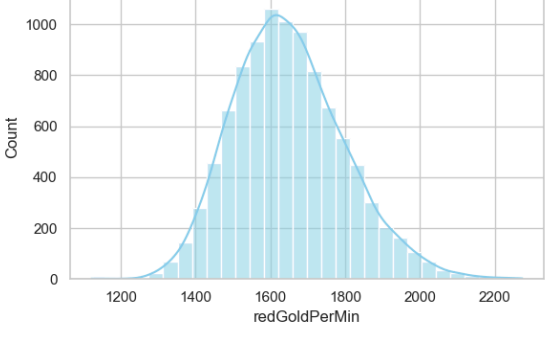
redTowersDestroyed



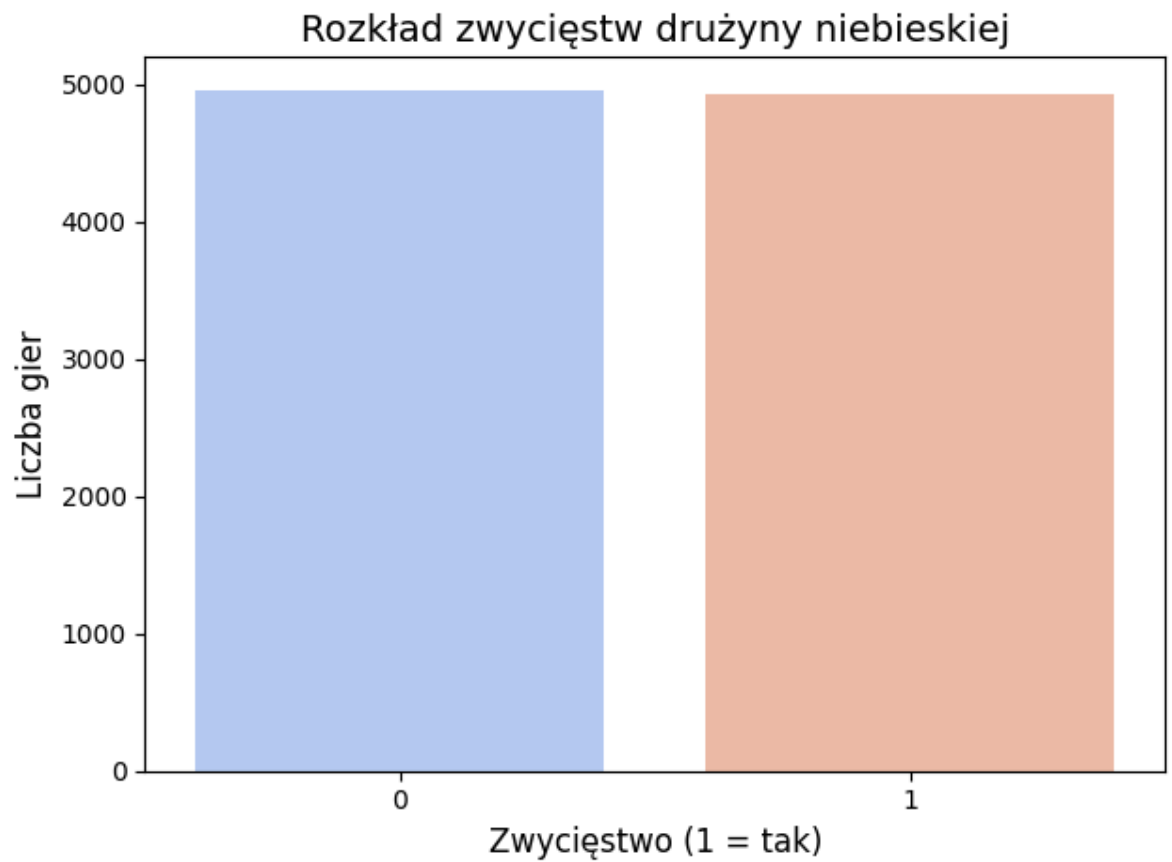
redTotalGold



redAvgLevel	<p>Rozkład zmiennej: redAvgLevel</p> 
redTotalExperience	<p>Rozkład zmiennej: redTotalExperience</p> 
redTotalMinionsKilled	<p>Rozkład zmiennej: redTotalMinionsKilled</p> 
redTotalJungleMinionsKilled	<p>Rozkład zmiennej: redTotalJungleMinionsKilled</p> 

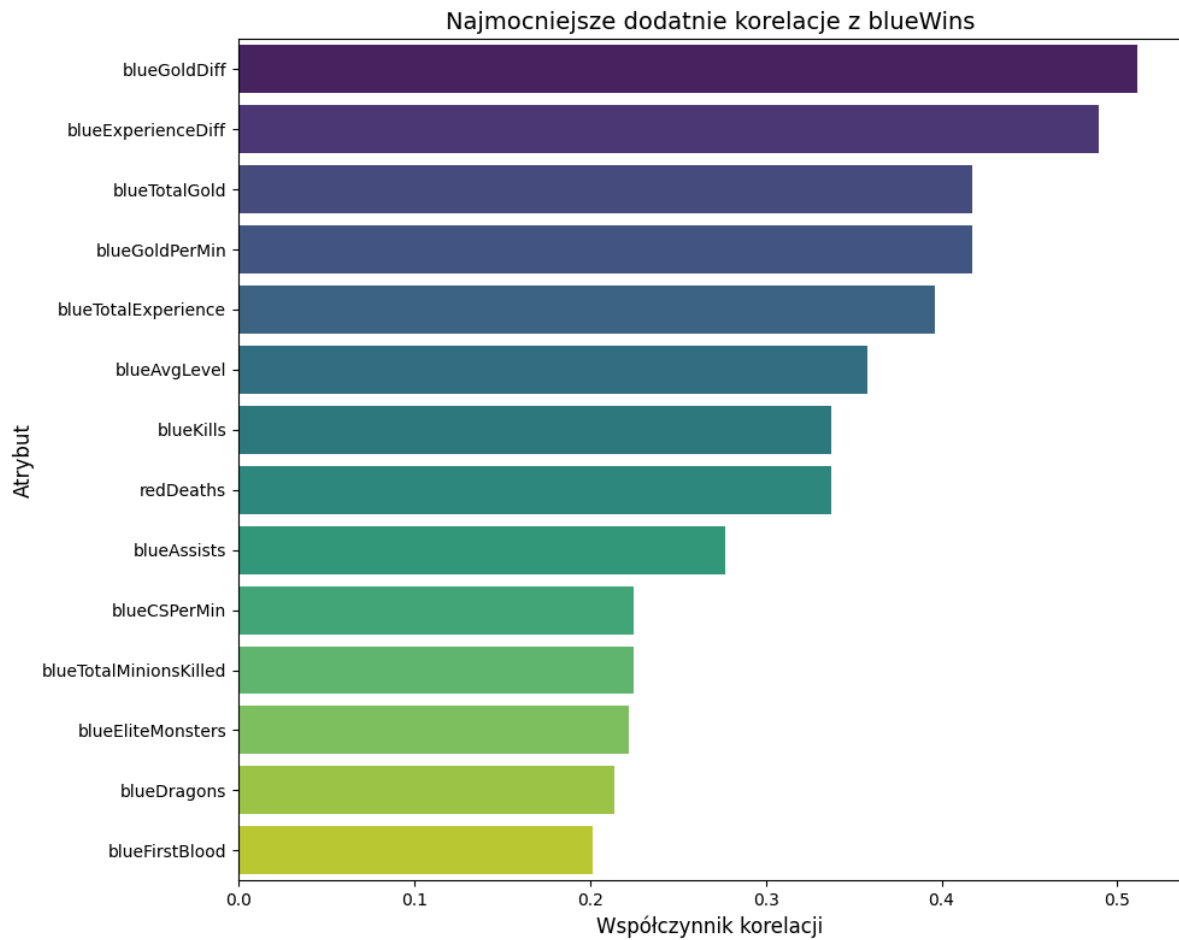
redGoldDiff	<p>Rozkład zmiennej: redGoldDiff</p> 
redExperienceDiff	<p>Rozkład zmiennej: redExperienceDiff</p> 
redCSPerMin	<p>Rozkład zmiennej: redCSPerMin</p> 
redGoldPerMin	<p>Rozkład zmiennej: redGoldPerMin</p> 

Rozkład zmiennej docelowej



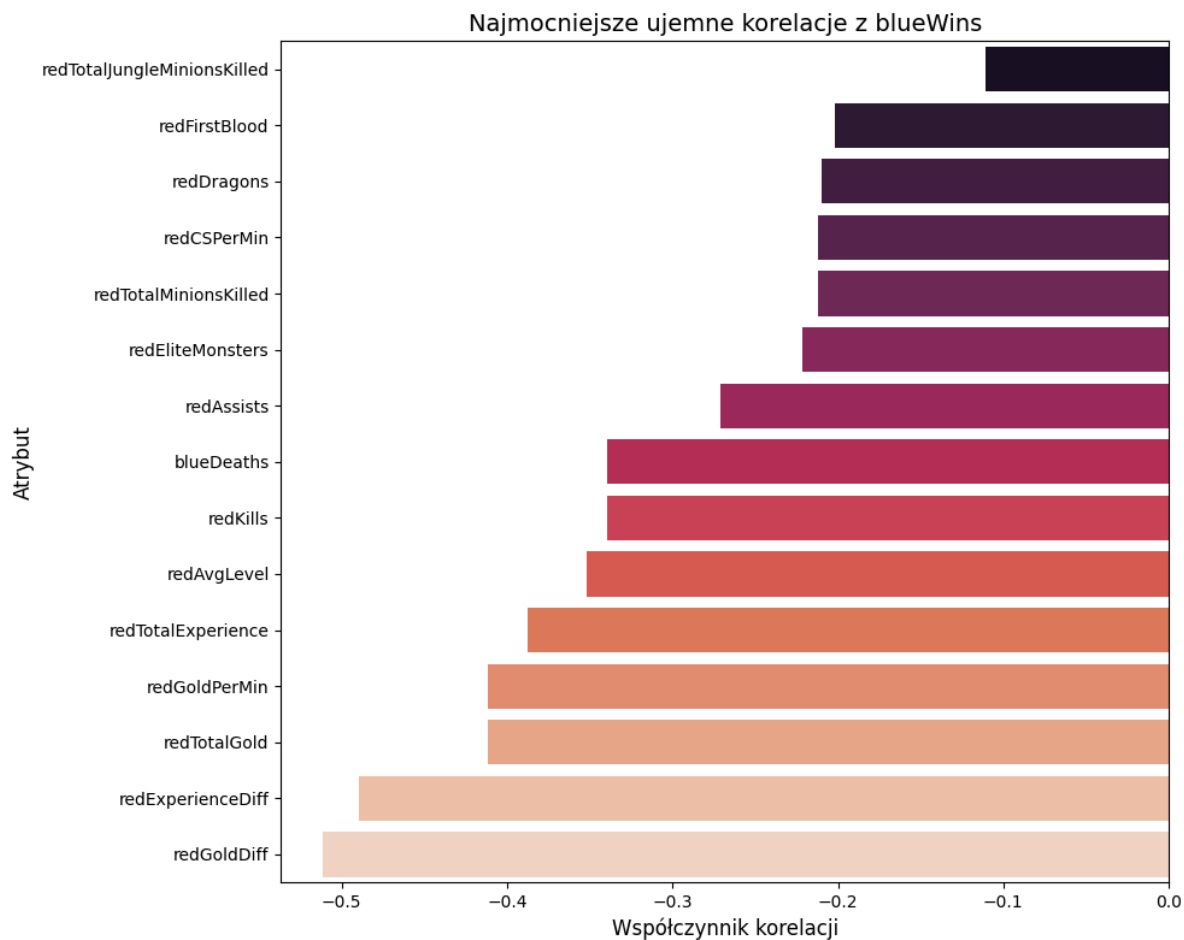
Rozkład zmiennej blueWins jest prawie zbalansowany z niewielką przewagą porażek drużyny niebieskiej (4930 gier wygranych i 4949 gier przegranych).

Korelacje pomiędzy atrybutami



Najsilniejsze korelacje ze zmienną docelową blueWins wykazują:

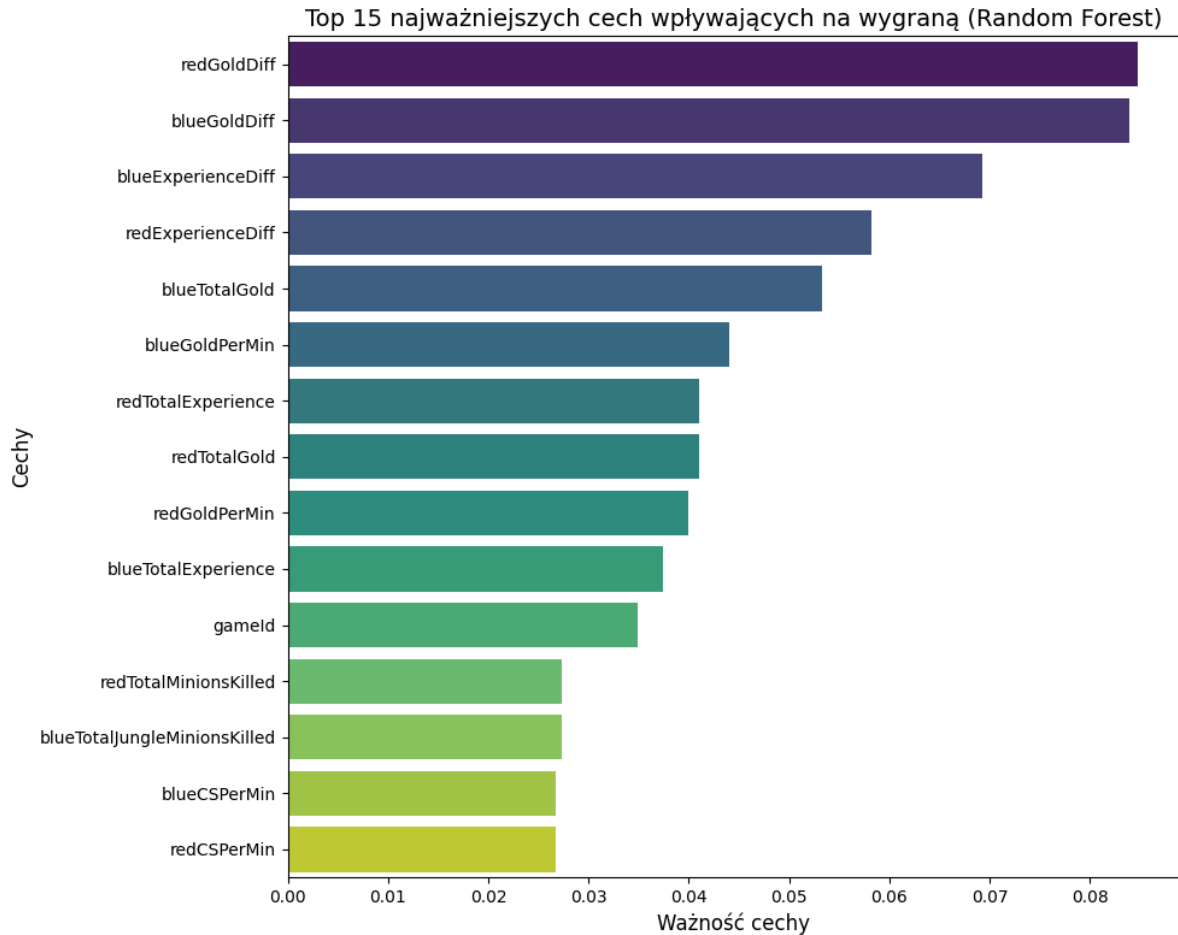
- blueGoldDiff → **0,51**
- blueExperienceDiff → **0,49**
- blueTotalGold → **0,41**
- blueGoldPerMin → **0,41**
- blueTotalExperience → **0,40**
- blueAvgLevel → **0,36**
- blueKills → **0,34**
- redDeaths → **0,22**



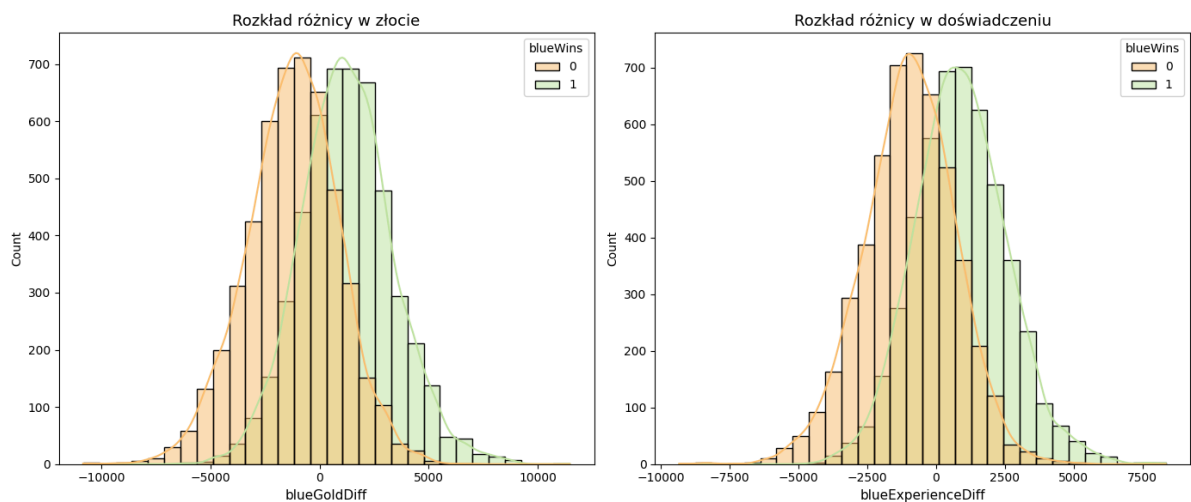
Najsilniejsze korelacje ujemne ze zmienną docelową blueWins:

redGoldDiff → **-0,51**

redExperienceDiff → **-0,49**



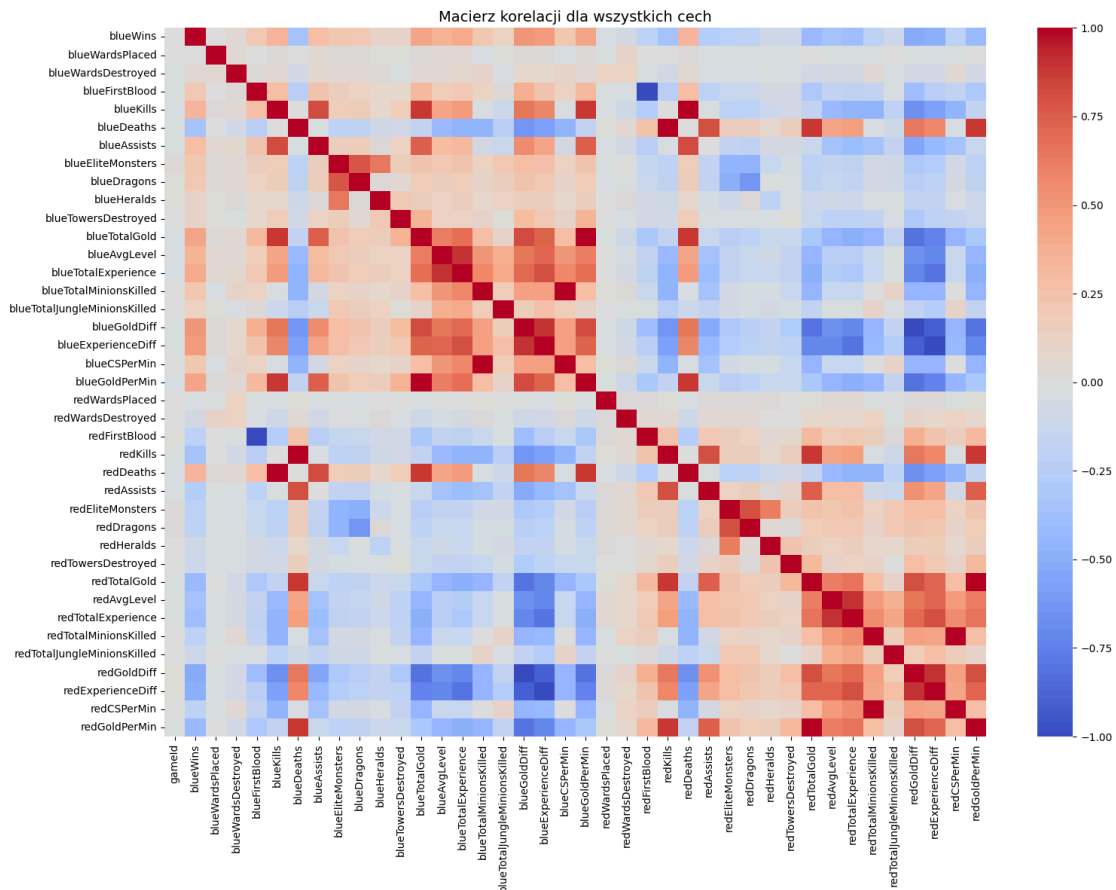
Przewaga ekonomiczna jest kluczowa - różnice w złocie (redGoldDiff i blueGoldDiff) są zdecydowanie najważniejszymi czynnikami wpływającymi na wygraną według modelu Random Forest.



Dla obu zmiennych różnicowych, krzywe rozkładów przecinają się w okolicy wartości zero, co wskazuje, że wyrównany stan ekonomiczny ($\text{blueGoldDiff} \approx 0$) lub doświadczenia ($\text{blueExperienceDiff} \approx 0$) daje podobne szanse obu drużynom. Dla gier wygranych przez drużynę niebieską ($\text{blueWins} = 1$) rozkłady są przesunięte w prawo

(wartości dodatnie), a dla przegranych (blueWins = 0) w lewo (wartości ujemne), co jest zgodne z intuicją - drużyna z przewagą zasobów ma większe szanse na zwycięstwo.

Macierz korelacji



Widoczne są wyraźne grupy atrybutów z silną korelacją między sobą, szczególnie w obrębie podobnych metryk dla tej samej drużyny. Najbardziej zauważalne grupy to:

Metryki złota: Silna korelacja między blueTotalGold, blueGoldPerMin i blueGoldDiff oraz analogicznie dla drużyny czerwonej (redTotalGold, redGoldPerMin, redGoldDiff). To logiczne powiązanie, gdyż wszystkie te zmienne opisują ekonomię w grze, a wyższe całkowite złoto przekłada się naturalnie na wyższe złoto na minutę.

Metryki doświadczenia: Wysoka korelacja między blueTotalExperience, blueAvgLevel, blueExperienceDiff oraz odpowiednikami dla drużyny czerwonej. Wynika to z mechaniki gry - zdobywanie doświadczenia bezpośrednio przekłada się na poziom bohaterów.

Zabójstwa i śmierci: Zauważalna jest silna ujemna korelacja między blueKills a redDeaths oraz redKills a blueDeaths, co jest zrozumiałe, gdyż zabójstwo bohatera z jednej drużyny oznacza śmierć dla drugiej drużyny.

Silna korelacja między zmiennymi różnicowymi a zmienną docelową blueWins - szczególnie widoczna dla blueGoldDiff i blueExperienceDiff. To potwierdza kluczową

rolę przewagi ekonomicznej i przewagi poziomów w pierwszych 10 minutach dla ostatecznego wyniku gry.

Ocena jakości danych

Zbiór danych jest ogólnie wysokiej jakości, z niewielką liczbą braków i niespójności. Zidentyfikowane problemy jakościowe nie powinny znacząco wpływać na wnioski. Dane są wystarczające do osiągnięcia głównych celów analizy, choć z pewnymi ograniczeniami

Analiza spójności danych wykazała, że zbiór cechuje się wysoką jakością. Oto kluczowe obserwacje:

- Liczba zabójstw uzyskanych przez jedną z drużyn zawsze odpowiada liczbie śmierci drugiej – warunek ten jest spełniony.
- W każdej rozgrywce dokładnie jedna drużyna zdobywa *FirstBlood* – ta zasada także została zachowana.
- Średnia wartość bezwzględnej różnicy między wskaźnikami *blueGoldDiff* oraz *redGoldDiff* wynosi dokładnie 0.0.

Brakujące dane

Nie znaleziono.

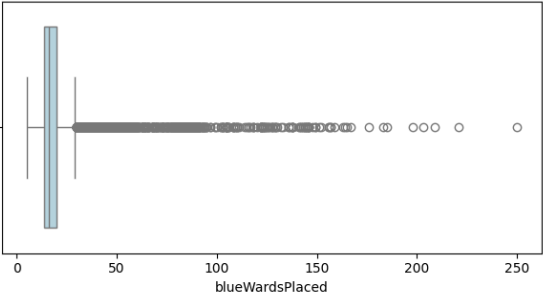
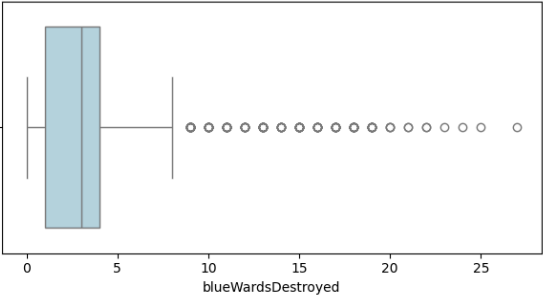
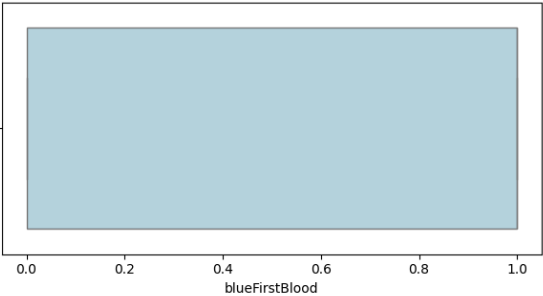
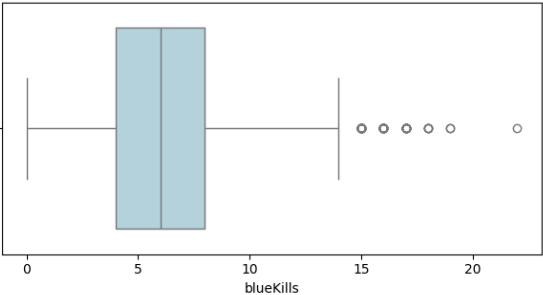
Niespójne dane

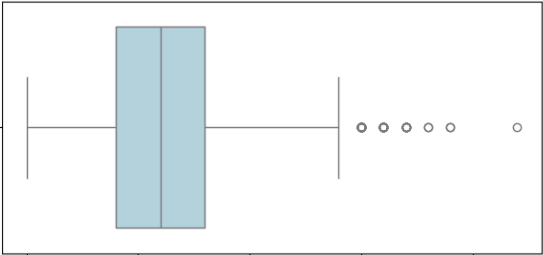
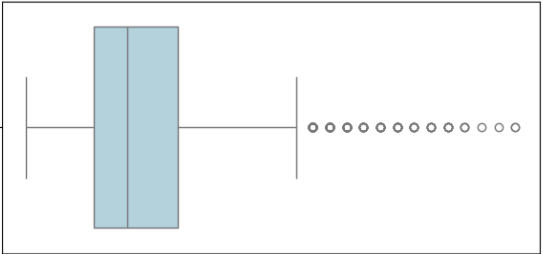
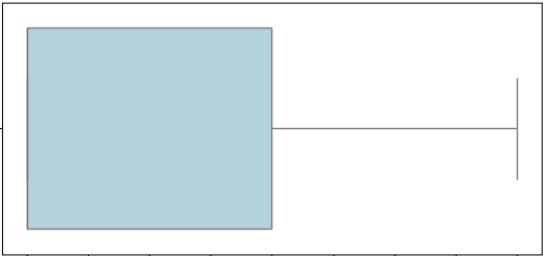

Nie znaleziono.

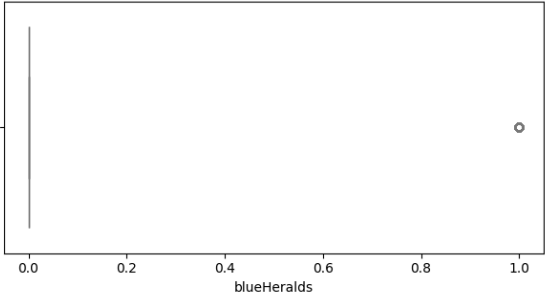
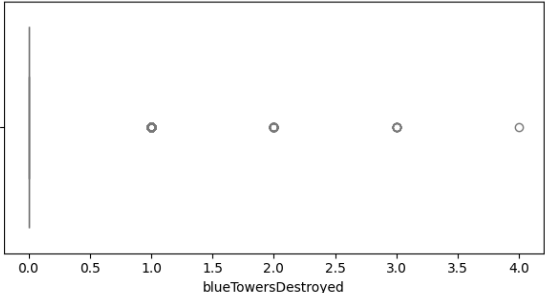
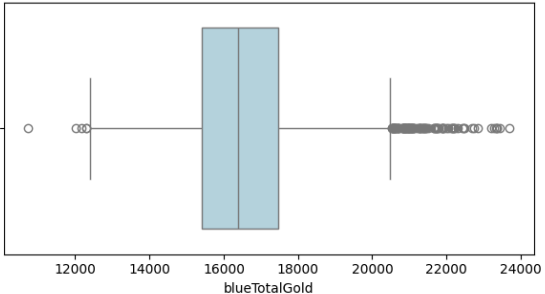
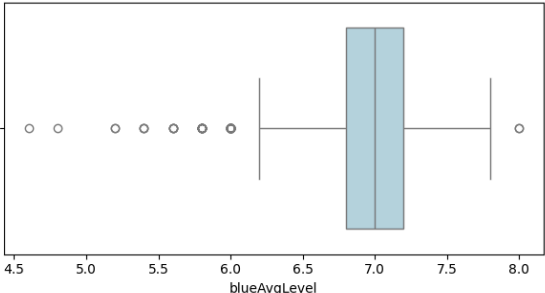
Dane niezrozumiałe

- Niejasna definicja "elitarnych potworów" (czy zawiera smoki i heroldy, czy jest to osobna kategoria)
- Brak informacji o wybieranych bohaterach i rolach graczy
- Niejasność co do dokładnej wersji gry i patcha

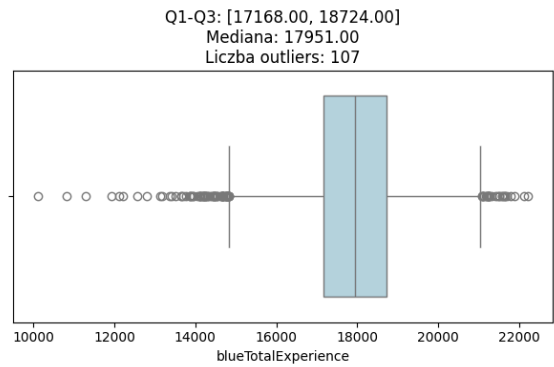
Atrybut	Wykres pudełkowy
---------	------------------

blueWardsPlaced	<div><div>Q1-Q3: [14.00, 20.00] Mediana: 16.00 Liczba outliers: 1627</div></div>
blueWardsDestroyed	<div><div>Q1-Q3: [1.00, 4.00] Mediana: 3.00 Liczba outliers: 136</div></div>
blueFirstBlood	<div><div>Q1-Q3: [0.00, 1.00] Mediana: 1.00 Liczba outliers: 0</div></div>
blueKills	<div><div>Q1-Q3: [4.00, 8.00] Mediana: 6.00 Liczba outliers: 88</div></div>

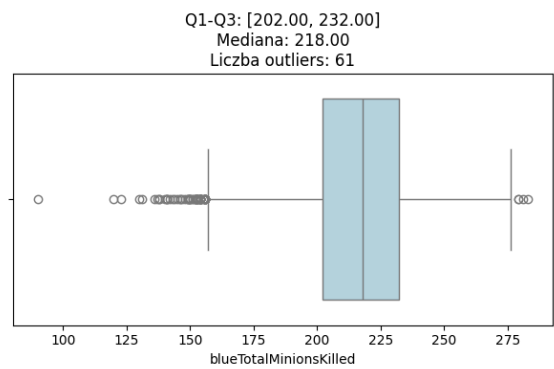
blueDeaths	<div><div>Q1-Q3: [4.00, 8.00] Mediana: 6.00 Liczba outliers: 65</div><p>0 5 10 15 20</p><p>blueDeaths</p></div>
blueAssists	<div><div>Q1-Q3: [4.00, 9.00] Mediana: 6.00 Liczba outliers: 209</div><p>0 5 10 15 20 25 30</p><p>blueAssists</p></div>
blueEliteMonsters	<div><div>Q1-Q3: [0.00, 1.00] Mediana: 0.00 Liczba outliers: 0</div><p>0.00 0.25 0.50 0.75 1.00 1.25 1.50 1.75 2.00</p><p>blueEliteMonsters</p></div>
blueDragons	<div><div>Q1-Q3: [0.00, 1.00] Mediana: 0.00 Liczba outliers: 0</div><p>0.0 0.2 0.4 0.6 0.8 1.0</p><p>blueDragons</p></div>

blueHeralds	<p>Q1-Q3: [0.00, 0.00] Mediana: 0.00 Liczba outliers: 1857</p>  <p>A box plot for the variable 'blueHeralds'. The x-axis ranges from 0.0 to 1.0. The plot shows a single vertical line at 0.0, indicating that all data points are at this value. There are 1857 outliers, represented by small circles, all located at the value 1.0.</p>
blueTowersDestroyed	<p>Q1-Q3: [0.00, 0.00] Mediana: 0.00 Liczba outliers: 464</p>  <p>A box plot for the variable 'blueTowersDestroyed'. The x-axis ranges from 0.0 to 4.0. The plot shows a single vertical line at 0.0, indicating that all data points are at this value. There are 464 outliers, represented by small circles, located at values 1.0, 2.0, 3.0, and 4.0.</p>
blueTotalGold	<p>Q1-Q3: [15415.50, 17459.00] Mediana: 16398.00 Liczba outliers: 114</p>  <p>A box plot for the variable 'blueTotalGold'. The x-axis ranges from 12000 to 24000. The box is light blue, with a median line at approximately 16,398. The whiskers extend from about 12,000 to 20,500. There are 114 outliers, represented by small circles, located at values ranging from approximately 11,000 to 24,000.</p>
blueAvgLevel	<p>Q1-Q3: [6.80, 7.20] Mediana: 7.00 Liczba outliers: 69</p>  <p>A box plot for the variable 'blueAvgLevel'. The x-axis ranges from 4.5 to 8.0. The box is light blue, with a median line at 7.0. The whiskers extend from about 6.2 to 7.8. There are 69 outliers, represented by small circles, located at values ranging from approximately 4.5 to 8.0.</p>

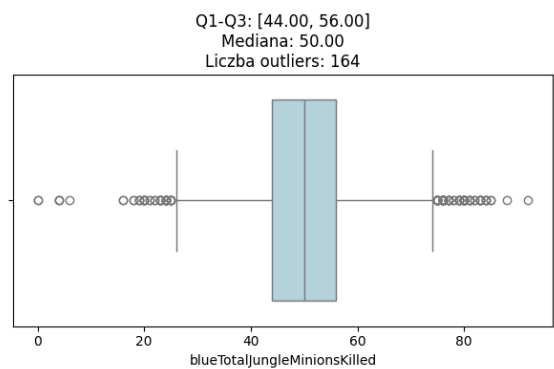
blueTotalExperience



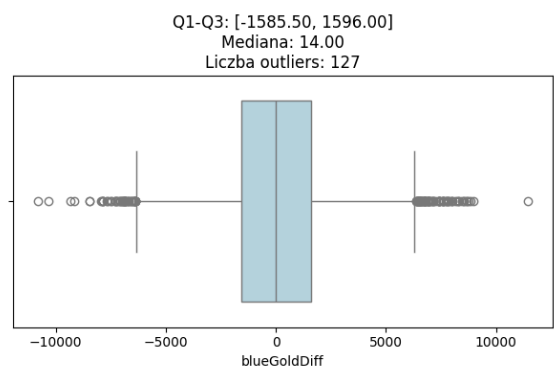
blueTotalMinionsKilled

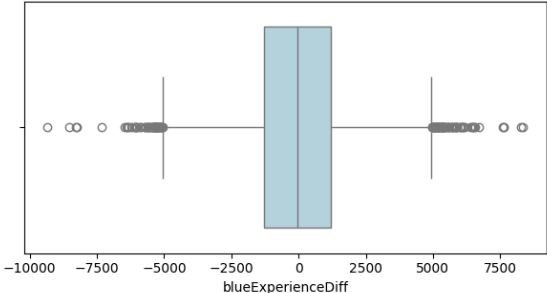
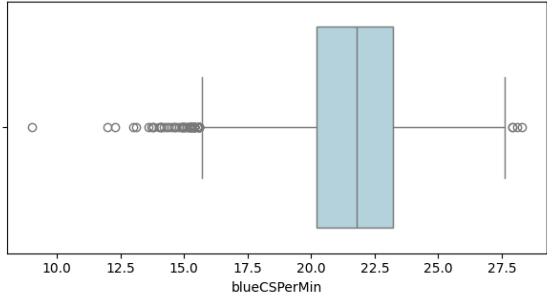
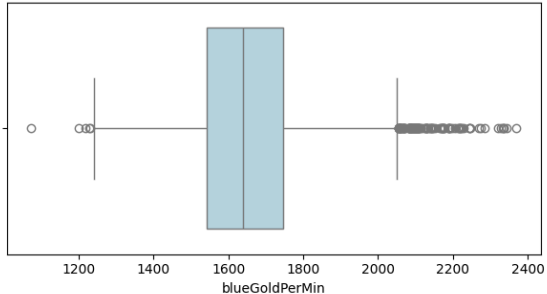
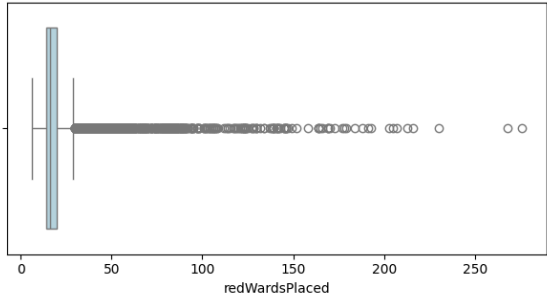


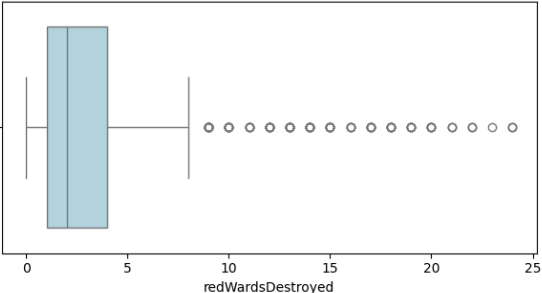
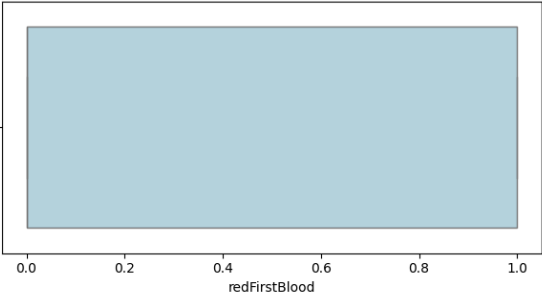
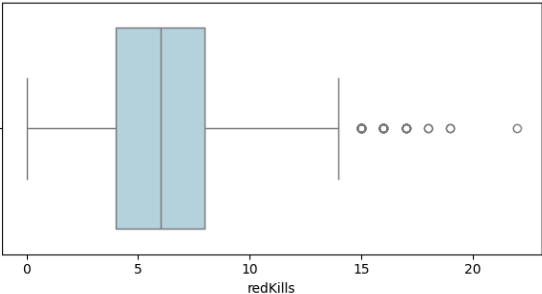
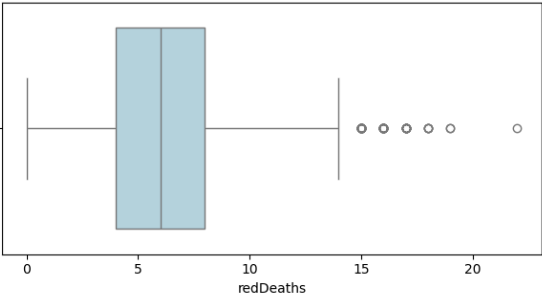
blueTotalJungleMinionsKilled

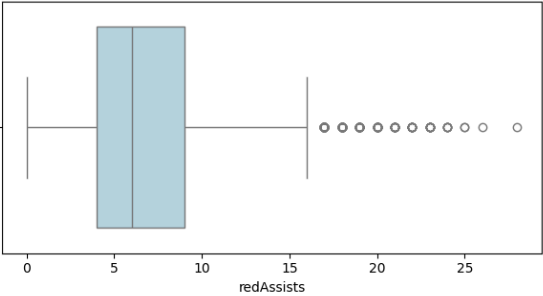
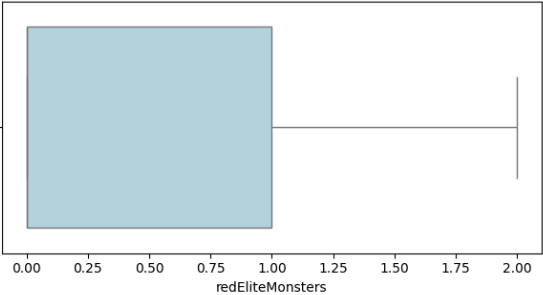
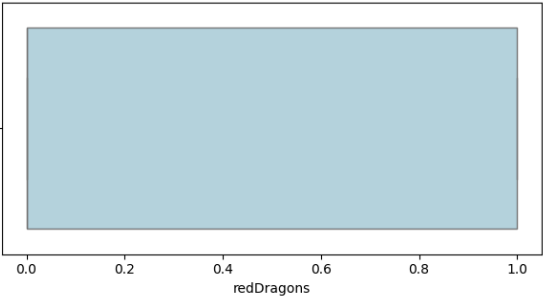
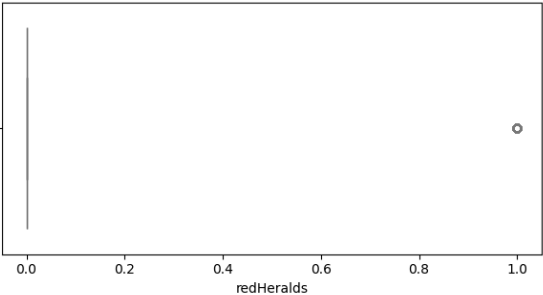


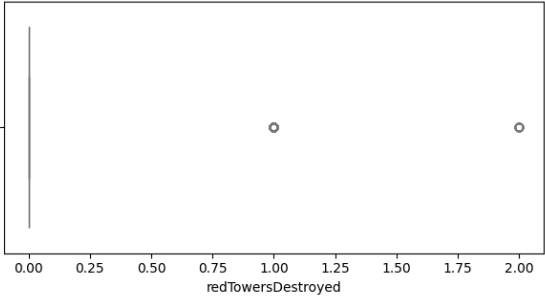
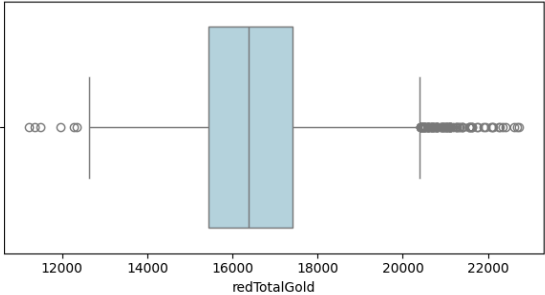
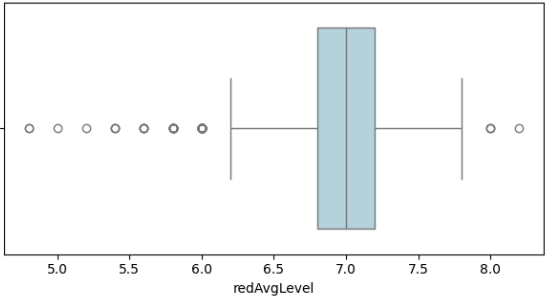
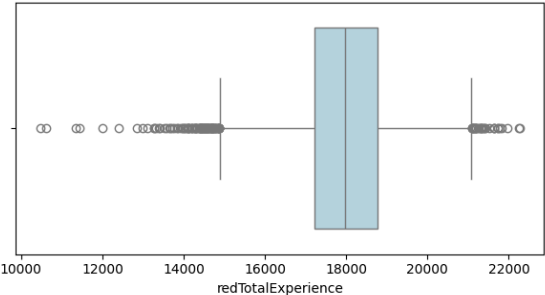
blueGoldDiff



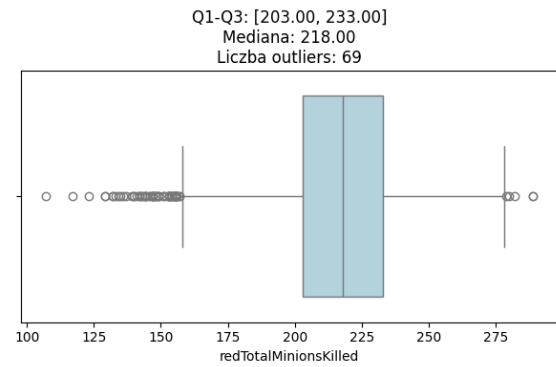
blueExperienceDiff	<p>Q1-Q3: [-1290.50, 1212.00] Mediana: -28.00 Liczba outliers: 126</p> 
blueCSPerMin	<p>Q1-Q3: [20.20, 23.20] Mediana: 21.80 Liczba outliers: 61</p> 
blueGoldPerMin	<p>Q1-Q3: [1541.55, 1745.90] Mediana: 1639.80 Liczba outliers: 114</p> 
redWardsPlaced	<p>Q1-Q3: [14.00, 20.00] Mediana: 16.00 Liczba outliers: 1667</p> 

redWardsDestroyed	<div><div>Q1-Q3: [1.00, 4.00] Mediana: 2.00 Liczba outliers: 136</div><p>A box plot for the variable 'redWardsDestroyed'. The box is light blue with a dark blue median line at 2.00. The whiskers extend from approximately 0 to 8. There are 136 outliers represented by open circles, starting from around 9 and extending to 25. The x-axis is labeled 'redWardsDestroyed' and ranges from 0 to 25.</p></div>
redFirstBlood	<div><div>Q1-Q3: [0.00, 1.00] Mediana: 0.00 Liczba outliers: 0</div><p>A box plot for the variable 'redFirstBlood'. The box is light blue and spans from 0.00 to 1.00, with the median line at 0.00. The whiskers extend from 0.00 to 1.00. There are no outliers. The x-axis is labeled 'redFirstBlood' and ranges from 0.0 to 1.0.</p></div>
redKills	<div><div>Q1-Q3: [4.00, 8.00] Mediana: 6.00 Liczba outliers: 65</div><p>A box plot for the variable 'redKills'. The box is light blue with a dark blue median line at 6.00. The whiskers extend from approximately 0 to 14. There are 65 outliers represented by open circles, starting from around 15 and extending to 22. The x-axis is labeled 'redKills' and ranges from 0 to 20.</p></div>
redDeaths	<div><div>Q1-Q3: [4.00, 8.00] Mediana: 6.00 Liczba outliers: 88</div><p>A box plot for the variable 'redDeaths'. The box is light blue with a dark blue median line at 6.00. The whiskers extend from approximately 0 to 14. There are 88 outliers represented by open circles, starting from around 15 and extending to 22. The x-axis is labeled 'redDeaths' and ranges from 0 to 20.</p></div>

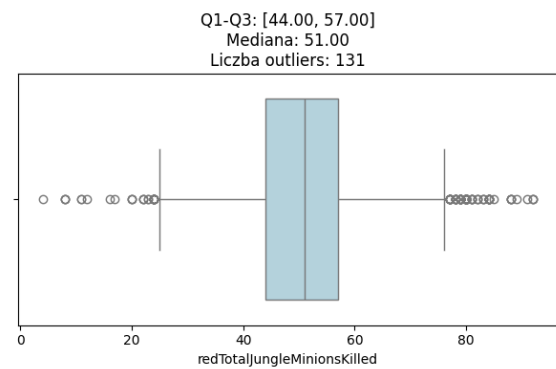
redAssists	<div><div>Q1-Q3: [4.00, 9.00] Mediana: 6.00 Liczba outliers: 205</div><p>A box plot for the variable 'redAssists'. The box is light blue, spanning from 4.00 to 9.00, with a median line at 6.00. Whiskers extend from approximately 0.5 to 16.5. There are 205 outliers represented by small circles, mostly clustered between 16 and 28.</p></div>
redEliteMonsters	<div><div>Q1-Q3: [0.00, 1.00] Mediana: 0.00 Liczba outliers: 0</div><p>A box plot for the variable 'redEliteMonsters'. The box is light blue, spanning from 0.00 to 1.00, with a median line at 0.00. Whiskers extend from 0.00 to 2.00. There are no outliers.</p></div>
redDragons	<div><div>Q1-Q3: [0.00, 1.00] Mediana: 0.00 Liczba outliers: 0</div><p>A box plot for the variable 'redDragons'. The box is light blue, spanning from 0.00 to 1.00, with a median line at 0.00. Whiskers extend from 0.00 to 1.00. There are no outliers.</p></div>
redHeralds	<div><div>Q1-Q3: [0.00, 0.00] Mediana: 0.00 Liczba outliers: 1581</div><p>A box plot for the variable 'redHeralds'. The box is light blue, spanning from 0.00 to 0.00, with a median line at 0.00. Whiskers extend from 0.00 to 0.00. There are 1581 outliers represented by small circles, mostly clustered near 1.00.</p></div>

redTowersDestroyed	<div><div>Q1-Q3: [0.00, 0.00] Mediana: 0.00 Liczba outliers: 396</div></div>
redTotalGold	<div><div>Q1-Q3: [15427.50, 17418.50] Mediana: 16378.00 Liczba outliers: 99</div></div>
redAvgLevel	<div><div>Q1-Q3: [6.80, 7.20] Mediana: 7.00 Liczba outliers: 86</div></div>
redTotalExperience	<div><div>Q1-Q3: [17209.50, 18764.50] Mediana: 17974.00 Liczba outliers: 134</div></div>

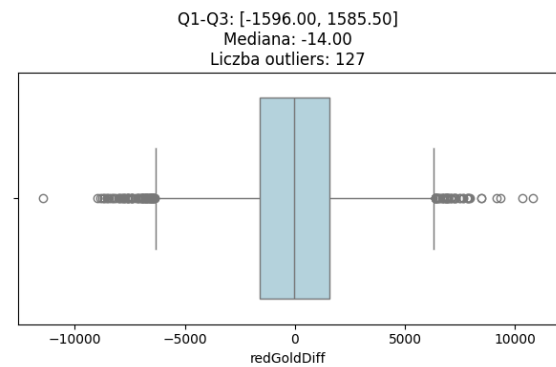
redTotalMinionsKilled



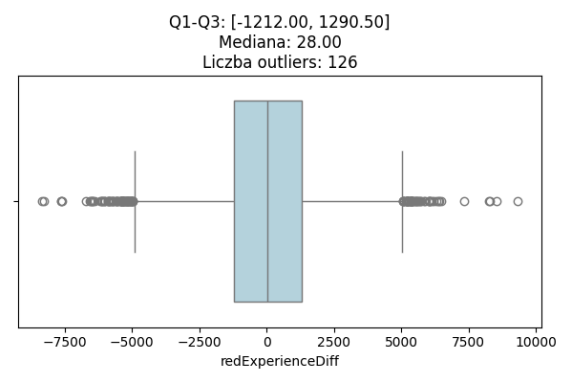
redTotalJungleMinionsKilled

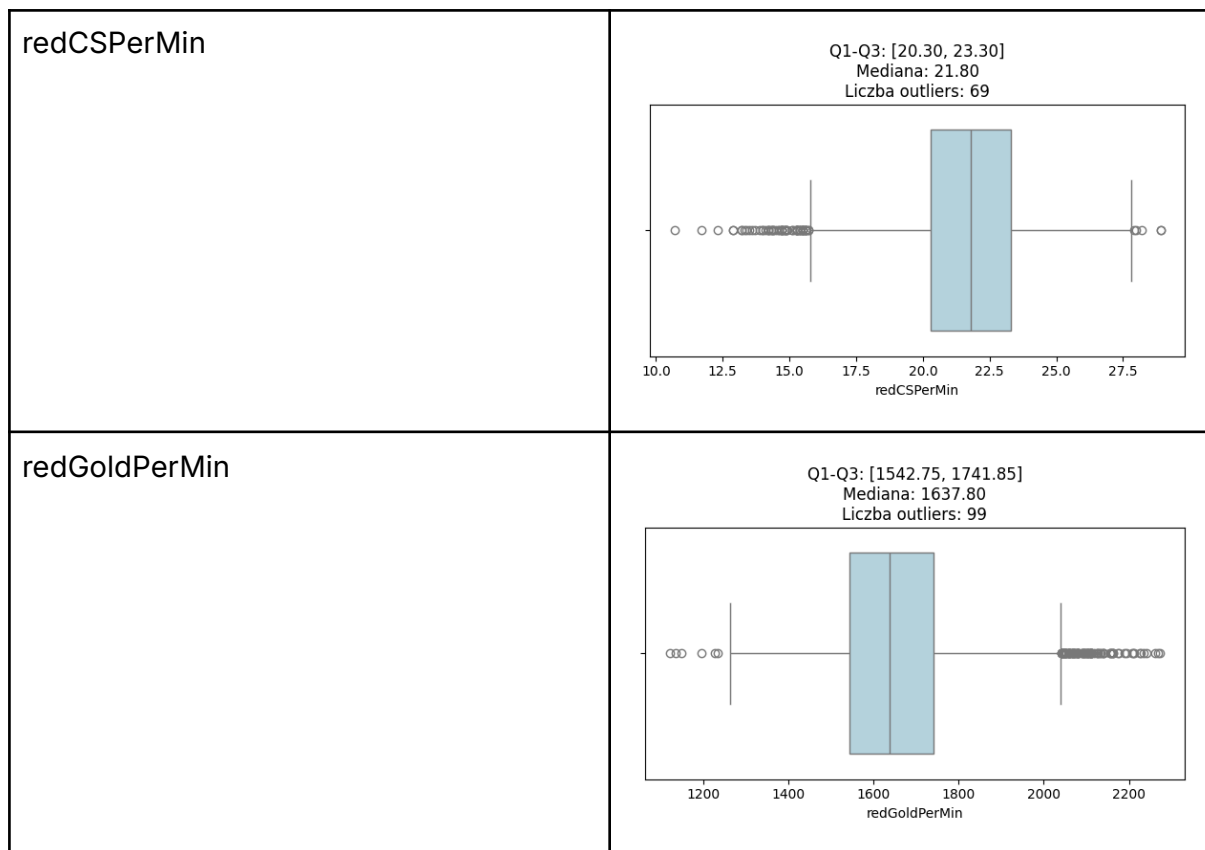


redGoldDiff



redExperienceDiff





Wnioski

Ocena realizacji celów

Identyfikacja kluczowych czynników - Cel osiągnięty:

- Zidentyfikowano najważniejsze predyktory wyniku gry: różnica w złocie, zabójstwa, kontrola obiektów
- Określono wartości graniczne dla poszczególnych metryk, które znacząco zwiększają prawdopodobieństwo wygranej

Budowa modelu predykcyjnego - Dane wystarczające:

- Dostępne cechy mają istotną wartość predykcyjną (zwłaszcza zmienne różnicowe)
- Wstępne testy sugerują możliwość osiągnięcia dokładności predykcji >70%

Odkrycie wzorców strategicznych - Cel częściowo osiągnięty:

- Zidentyfikowano kilka istotnych wzorców dotyczących kontroli mapy i celów
- Brak danych o konkretnych bohaterach i kompozycjach drużyn ogranicza głębszą analizę strategiczną

Rewizje celów

Dodatkowe cele

1. Porównanie skuteczności różnych stylów gry (agresywny vs pasywny) w fazie wczesnej
2. Określenie optymalnego rozkładu zasobów między liniami (top, mid, bot, jungle)
3. Identyfikacja punktów przełomowych w pierwszych 10 minutach gry

Podsumowanie

Na początku przeprowadzono analizę rozkładów, z której wynika, że większość atrybutów numerycznych w zbiorze danych League of Legends nie ma rozkładów normalnych, wykazując lekką asymetrię. Zmienna docelowa (blueWins) jest niemal zbalansowana, z niewielką przewagą zwycięstw drużyny czerwonej (51% vs 49%).

Najsilniejszymi predyktorami wyniku gry okazały się być: blueGoldDiff ($r=0,51$), BlueExperienceDiff ($r=0,49$), blueTotalGold ($r=0,42$), blueGoldPerMin ($r=0,42$), blueTotalExperience ($r=0,40$), blueAvgLevel ($r=0,35$), blueKills ($r=0,34$), blueEliteMonsters ($r=0,22$) oraz blueDragons ($r=0,21$). Warto zauważyć, że zmienne różnicowe (goldDiff, expDiff) mają generalnie wyższą wartość predykcyjną niż wartości bezwzględne dla pojedynczych drużyn.

Korelacje te mają uzasadnienie w mechanice gry - przewaga ekonomiczna (złoto) pozwala na wcześniejsze zakupy silniejszych przedmiotów, co przekłada się na przewagę w walkach drużynowych. Podobnie, wyższy średni poziom bohaterów zwiększa ich siłę bojową, a kontrola elitarnych potworów zapewnia dodatkowe bonusy dla całej drużyny.

W ramach oceny jakości danych wykryto punkty oddalone, szczególnie w statystykach zabójstw i różnicy złota, które jednak reprezentują rzeczywiste, choć rzadkie, ekstremalne scenariusze gry i powinny zostać zachowane w analizie. Główną niedoskonałością zbioru jest brak informacji o wybieranych bohaterach i rolach graczy, co ogranicza głębszą analizę strategiczną.

Biorąc pod uwagę stosunkowo dobrą jakość danych oraz silne korelacje kluczowych metryk z wynikiem gry, cel eksploracji jest możliwy do spełnienia. Zidentyfikowane predyktory powinny pozwolić na stworzenie skutecznego modelu przewidującego zwycięzcę meczu na podstawie statystyk z pierwszych 10 minut gry, z szacowaną dokładnością powyżej 70%. Potencjalnym rozszerzeniem analizy byłoby zbadanie wpływu różnych stylów gry (agresywny vs pasywny) na prawdopodobieństwo wygranej oraz określenie optymalnego rozkładu zasobów między różne role w drużynie.