

Topologiczna analiza danych w tenisie stołowym na podstawie 1. ligi kobiet.

Aleksandra Garb

1. Dane.

Do projektu wykorzystałam dane z rundy zasadniczej 1. ligi kobiet grupy południowej w tenisie stołowym w sezonie 2024/2025. Informacje te zostały udostępnione na stronie Polskiego Związku Tenisa Stołowego, a także uzupełniłam je na podstawie własnej wiedzy.

W analizie statystycznej uwzględniłam po sześć najlepszych zawodniczek z każdej z dziesięciu drużyn ligi. W celu stworzenia punktacji, dla każdej zawodniczki obliczyłam szereg parametrów, takich jak:

- wiek,
- skuteczność,
- liczba punktów zdobytych i straconych,
- styl gry,
- procent meczów wygranych na własnej hali,
- procent meczów wygranych na wyjeździe,
- procent zwycięstw z zawodniczkami grającymi długimi czopami oraz stylem defensywnym,
- procent wygranych piątych setów,
- procent setów wygranych różnicą dwóch punktów,
- procent wygranych gier podwójnych.

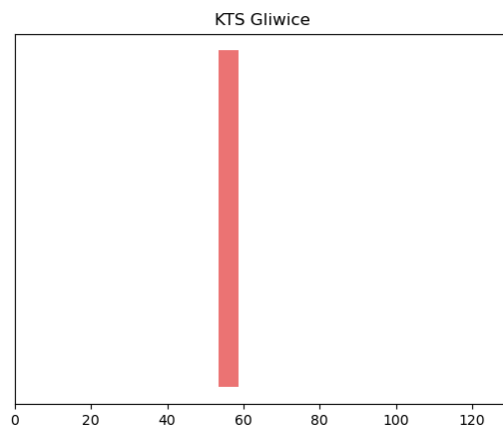
W przypadku braku możliwości określenia któregoś z parametrów na podstawie dostępnych danych, przyjmowałam wartość środkową, czyli 50%.

2. Opis analizy.

W swojej analizie postanowiłam zbadać obecność dziur zerowymiarowych, jednowymiarowych i dwuwymiarowych. Już po wykonaniu filtracji można było łatwo zauważyć, że nie pojawiły się żadne dziury dwuwymiarowe, dlatego ten przypadek został od razu wykluczony z dalszych rozważań.

W przypadku dwóch drużyn zaobserwowano występowanie dziur jednowymiarowych. Jednak drużyny te nie znajdowały się blisko siebie w tabeli

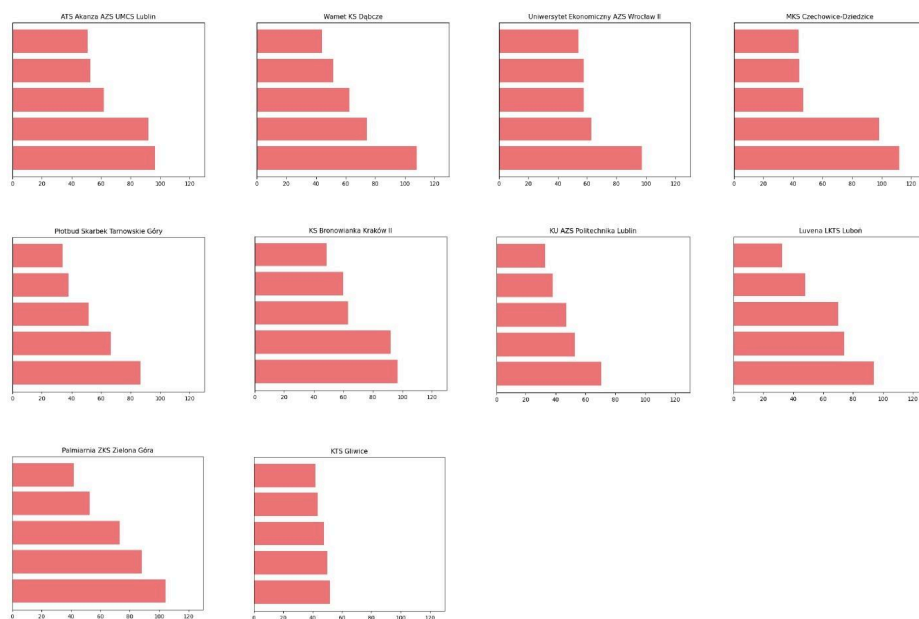
ligowej, a same dziury były bardzo krótkie. Z dużym prawdopodobieństwem powstały one przypadkowo i stanowiły tzw. "szum", dlatego również hipoteza o ich istotnym wpływie na analizę została szybko odrzucona.



Rys. 1 Przykład powstałej dziury jednowymiarowej.

3. Analiza dziur zerowymiarowych.

Pierwszym krokiem w analizie dziur zerowymiarowych było uporządkowanie wykresów wszystkich drużyn zgodnie z miejscami, jakie zajęły w tabeli po rundzie zasadniczej.



Rys. 2 Wykresy drużyn ustawione w kolejności miejsc w tabeli.

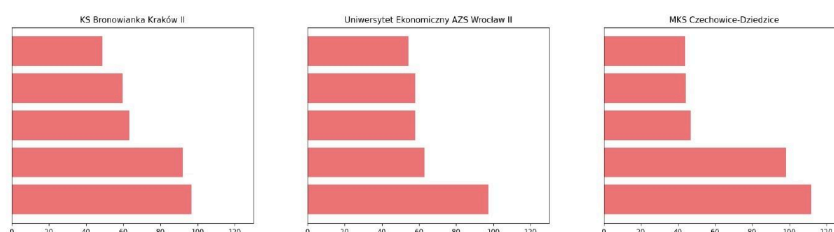
Niestety, analiza ta nie przyniosła jednoznacznych wniosków. Kształt trójkąta przyjmują wykresy zarówno drugiej, jak i dziewiątej drużyny. Ciekawą obserwację można natomiast wyciągnąć, porównując wykresy dziewiątej i dziesiątej drużyny – czyli dwóch ostatnich zespołów ligi. Wykres Zielonej Góry przyjmuje kształt trójkąta, natomiast wykres Gliwic jest bardzo równy. Najprawdopodobniej wynika to z faktu, że w drużynie z Gliwic grało sześć zawodniczek o bardzo zbliżonym poziomie – niestety niskim, ponieważ wszystkie miały słabe statystyki, a w Zielonej Górze zawodniczki zróżnicowane, ale nie na tyle dobre, żeby móc przebić się wyżej w lidze.

Jeśli chodzi o czołówkę tabeli, jedynym zauważalnym wnioskiem jest to, że aby wygrać ligę, drużyna musi mieć dwie bardzo mocne zawodniczki, jedną nieco słabszą, ale niewyraźnie odstającą, natomiast czwarta zawodniczka nie odgrywa już tak istotnej roli – jej statystyki mogą być znacznie słabsze.

4. Analiza modelu szkoleniowego klubów.

Drugim krokiem analizy była ocena sposobu szkolenia prowadzonego przez kluby uczestniczące w rozgrywkach. Można tutaj zauważyć, że drużyny z klubów posiadających ośrodki szkoleniowe składają się z bardziej wyrównanych zawodniczek. Wynika to prawdopodobnie z faktu, że kluby te tworzą zgrane grupy treningowe, w których wszystkie zawodniczki trenują razem na podobnym poziomie.

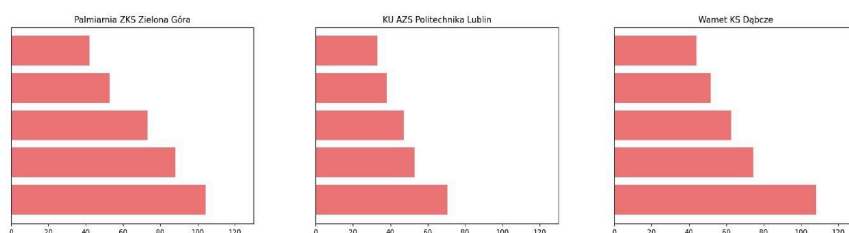
W szczególności wykresy drużyn z Wrocławia, Czechowic-Dziedzic oraz Krakowa pokazują, że zespoły te składały się z 4–5 zawodniczek o bardzo zbliżonych umiejętnościach, które stanowiły trzon drużyny. Widać tutaj model wspólnego szkolenia – ten sam trener, te same treningi, a zawodniczki te występowały w zdecydowanej większości meczów.



Rys. 3 Wykresy drużyn przy ośrodkach szkolenia.

Inny charakter mają drużyny, w których młode zawodniczki przeplatają się z bardziej doświadczonymi, ściągniętymi jedynie na potrzeby gry ligowej. Większość z tych zawodniczek trenuje w swoich macierzystych klubach lub w innych miastach, pracując z różnymi trenerami i reprezentując różne style gry. Wyjątkiem jest tu

drużyna KTS Gliwice, która jednak – z powodu bardzo słabych statystyk wszystkich zawodniczek i wyraźnego spadku z ligi – została pominięta w dalszej analizie.



Rys. 4 Przykład wykresów drużyn z zawodniczkami kupowanymi do ligi.

5. Klastrowanie.

Ostatnim etapem analizy było klastrowanie. Na początku stosowałam metodę klastrowania hierarchicznego. Niestety, ani grupowanie na dwie, ani na trzy klastry nie przyniosło oczekiwanych rezultatów – nie udało się znaleźć rzeczywistych powiązań pomiędzy drużynami przyporządkowanymi do poszczególnych klastrów.

W kolejnym podejściu zmieniłam metrykę – zamiast odległości Wassersteina zastosowałam odległość bottleneck. W tym przypadku do trzeciego klastra trafiła wyłącznie drużyna z Gliwic, co można uznać za sukces, ponieważ – jak wcześniej wspomniano – wyraźnie wyróżniała się na tle pozostałych zespołów.

Ciekawym wynikiem było również to, że do drugiego klastra zaklasyfikowano Politechnikę Lubelską i Skarbak Tarnowskie Góry – jedyne zespoły z „środką tabeli”, które nie są związane z ośrodkami szkoleniowymi. Może to być znacząca obserwacja, jednak ze względu na ograniczoną liczbę danych nie należy wyciągać pochopnych wniosków.

Na zakończenie spróbowałam jeszcze zastosować metrykę sliced Wasserstein, jednak i to podejście nie przyniosło oczekiwanych rezultatów.

6. Końcowe wnioski.

Topologiczna analiza danych z pierwszej ligi kobiet tenisa stołowego (grupa południowa) w Polsce w sezonie 2024/2025 nie przyniosła oczekiwanych rezultatów w kontekście przewidywania siły drużyn czy ich miejsc w tabeli. Na podstawie wygenerowanych barkodów nie da się jednoznacznie określić potencjału zespołu ani trafnie przewidzieć jego pozycji na koniec sezonu.

Analiza pozwoliła jednak na sformułowanie kilku innych, ciekawych wniosków. W Polsce funkcjonuje kilka klubów pełniących rolę ośrodków szkoleniowych, a zespoły z tych ośrodków – występujące na zapleczu Ekstraklasy – charakteryzują się wyrównanym poziomem zawodniczek. Kierownicy tych ośrodków konsekwentnie dążą do tego, by ich zawodniczki miały możliwość regularnej gry przez cały sezon, występując w podstawowym składzie. W efekcie drużyny te prezentują spójną strukturę i stabilność sportową.

Można również zauważyć pewną strategię budowania drużyn. Aby wygrać ligę, kluczowe jest posiadanie dwóch bardzo mocnych zawodniczek oraz trzeciej, nieco słabszej, ale wciąż na dobrym poziomie. Czwarta zawodniczka nie odgrywa już tak istotnej roli – jej słabsze statystyki nie wpływają znacząco na końcowy wynik zespołu. Z kolei, by utrzymać się w lidze, wystarczające jest posiadanie trzech solidnych zawodniczek. W obu przypadkach rola czwartej zawodniczki pozostaje marginalna z punktu widzenia siły drużyny.