What Kinds of Doubles Volleyball Partnerships are Successful?

An exercise in clustering

Spencer Matthews
University of California, Irvine

Abbreviated Abstract: Conventional beach volleyball knowledge says that a good team is typically made up of one "big" player and one all-around player. The job of the big is to block and hit, while the all-around player focuses on defense and setting (while still being able to hit a ball if needed). Here we analyze data from men's professional beach volleyball leagues in order to classify players as one of 4 groups (using k-means clustering) and determine which combinations of player groups are most successful. We find that for two of the player groups, success comes more often when they play with a partner from the same group, while the other two groups are more successful when paired with a player from a different group. We explore various metrics in the most successful pairing and discuss future applications using similar methods.

spencer.matthews@uci.edu

UCI University of California, Irvine

UCSAS 2021

Problem, Data, and Initial Clustering

Problem:
- Sand volleyball consists of partner pairs competing against each other
- We want to use clustering in a novel way to examine relationships between partner group combination and success

Data:
- Obtained from Kaggle
- Contains data from professional men's sand volleyball leagues from around the world

Initial Clustering:
- Used k-means clustering to identify 4 clusters of players (based on where elbow occurred in within-cluster SS plot)
- Relationships between clustering variables can be seen in the plot to the right (clusters shown by color)
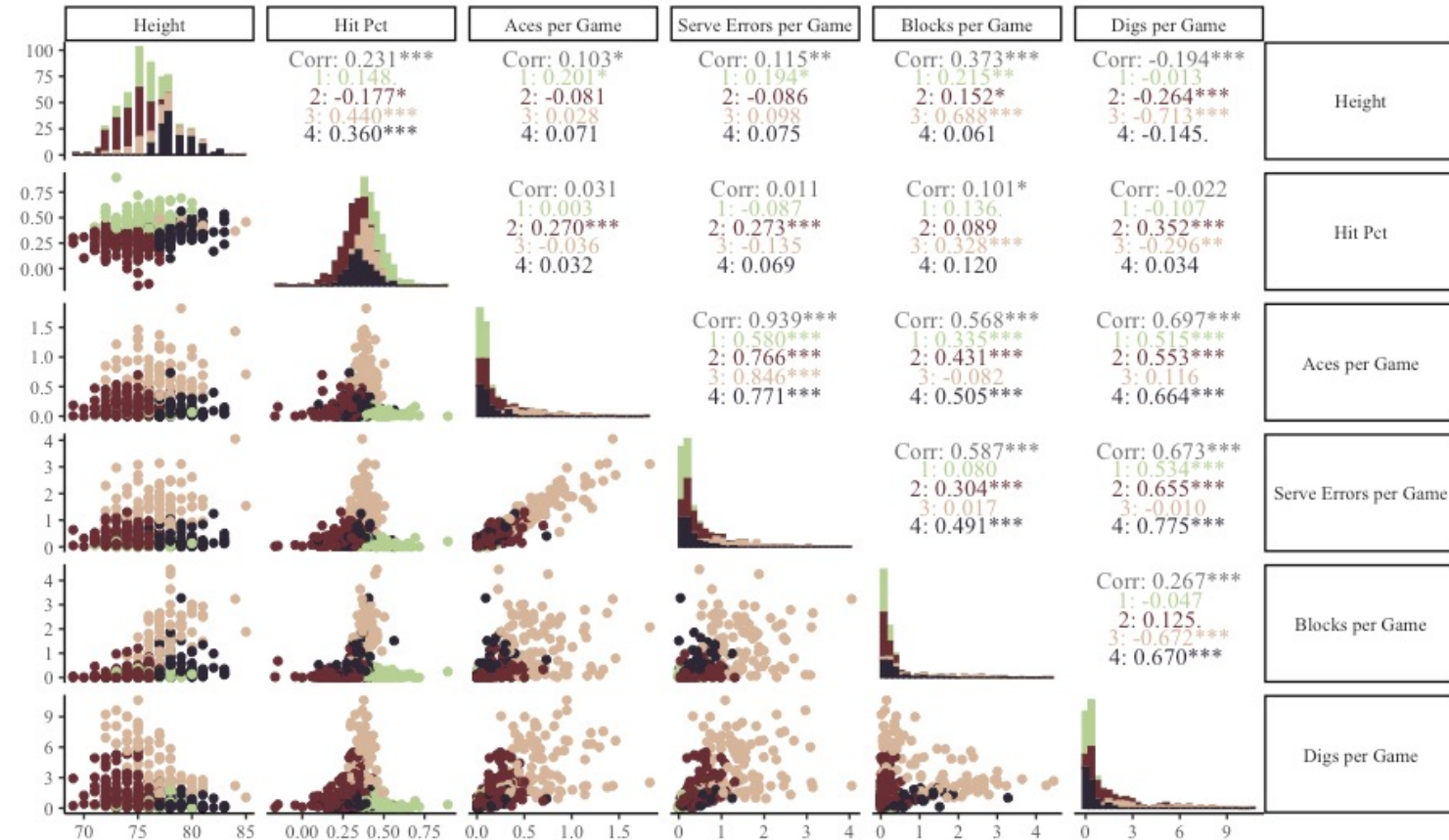- Cluster results were joined with game data to summarize by teammate combination

Figure 1: Variables used to cluster the players included height, hit percentage, aces per game, serve errors per game, blocks per game, and digs per game

spencer.matthews@uci.edu
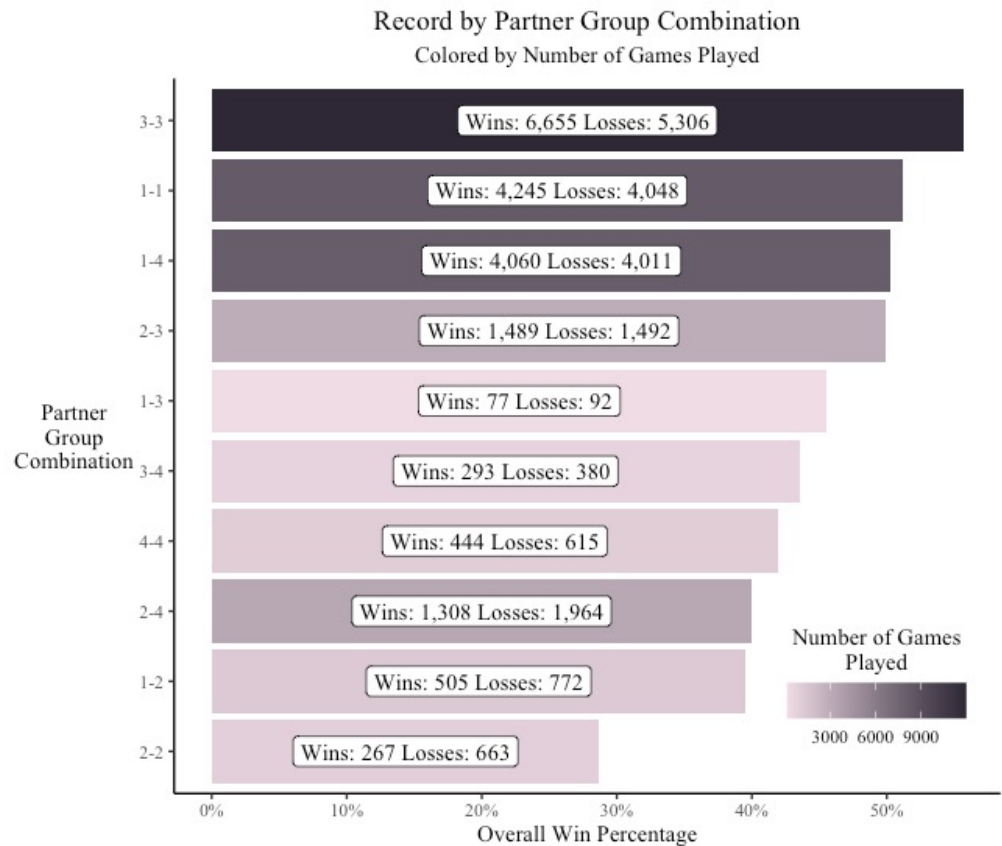
Analysis of Partner Group Combinations

Figure 2: Partner Group Combination results, with win percentage on the x axis and the combination on the y axis, darker bar indicates more total games in the dataset (note that the "better" combinations generally have a darker bar, which we would expect)

- The results in Figure 2 show that while some player groups seem to be better across the board (3), the top four partner group combinations contain players from all four groups.
- Also interesting is that while 3 and 1 play well with teammates from their own group, they do not play as well together (and hardly ever play together, for that matter)
- In Table 1 we see the results of the 3-3 partner group against all other partner groups
- The closest games on average are against 1-4 if judging by point differential, and 1-3 if judging by the match going to three games

Opponent Combination	Total Games	Record	Average Score Diff.	% Games that were a Sweep
1-4	783	53.51%	0.916	60.92%
1-1	691	56.73%	1.260	59.91%
1-3	56	55.36%	1.321	58.92%
2-3	1,529	59.71%	2.173	67.95%
1-2	141	74.46%	4.688	73.76%
3-4	311	73.31%	4.881	66.23%
4-4	125	71.20%	5.168	67.20%
2-4	611	78.72%	6.905	73.49%
2-2	338	90.08%	9.947	75.15%

Table 1: Partner group 3-3 stats against all other groups

spencer.matthews@uci.edu

UCSAS 2021

University of California, Irvine

Application and Next Steps

Potential Applications:
- Given a beach volleyball player, determine the optimal partner group
- Given an opponent, determine the partner group combination with the best possible chance of winning against the opponent
- As an organizer of a tournament, decide which match-ups are most likely to be close so that those can be televised or scheduled for "prime time"

Potential Problems/Limitations:
- Although the number of players in each cluster was about the same, the number of games played by the players in each cluster was very skewed (with 3s and 1s playing many more games than 2s and 4s) which could lead to bias in the results
- Although the initial problem talked about "big" players vs. all-around players, variables other than height were included in the analysis so the resulting groups are not clearly defined in terms of volleyball jargon

Next Steps:
- Perform a similar analysis for women's volleyball and compare the results – Are successful groupings in men's volleyball similar to successful groupings in women's volleyball?
- Analyze the cluster results by country/nationality – Are players from one country more likely to be in one cluster than another? Does a combination from one country significantly outperform the same combination from another country on average?
- Create an R shiny app so interested parties can explore the results

Additional Resources:
- View the github repo for a step-by-step walk-through of the analysis, including plot creation (https://github.com/srmatth/volleyball)
- For more information about the data, visit the Kaggle posting (https://www.kaggle.com/jessemostipak/beach-volleyball)