Liam Rosengren and Malcolm Gaynor
Olympics Project Executive Summary
Multivariate Sports Analytics | Spring 2024

To Lift Big or To Not Lift Big: A Case Study in Olympic Weightlifting Strategy

As student-athletes, we have a vested interest in lifting weights. After reading various literature, notably “Bayesian modeling of elite sporting performance with large databases” by Jim E. Griffin, Laurenţiu C. Hinoveanu and James G. Hopker, our objective became to analyze competitive weightlifting strategies at the olympic level.

By focusing solely on USA weightlifter Caine Wilkes, we were able to take a closer look at a multi-time world champion. Our data for this project came from the International Weightlifting Results Project (IWRP), which gave us all of Wilkes’ competition data dating back to 2004. We sought to find whether there were trends in the weights he chose for each iteration of his snatch and clean and jerk in competition. However, as the clean and jerk is the second lift in competition, we thought it would be more interesting to see if this lift yielded any observable strategies.

To begin, an exploratory data analysis found that Wilkes almost always completed his first attempt (28:3 success:failure), with a similar percentage on his second. However, his third attempt reduced drastically to a ratio of 10:21. This presents an opportunity for a strategic analysis of Wilkes’ attempt-wise weight decisions.

We followed this with a logistic regression model to predict the success rate of the third attempted clean and jerk (CJ3). A model that included each of the weights and a binary variable of success/failure for attempts one and two found significance at or above the 99% level for both age (negative coefficient) and bodyweight (positive). Intuitively, we thought it was sensible that a competitive weightlifter’s rate of success on their third attempt would decrease as they aged. The third attempt is typically the heaviest, either to claim a personal best or to win a competition, and these weights extract a heavy toll on the lifter. Additionally, the rate of success increasing as bodyweight increases represents a trend among the Super Heavyweight (109+ Kg) weight class.

Next, we constructed a few 4 state Markov Chains to model potential strategies for Caine Wilkes during the clean and jerk, using a decision tree to create the probabilities of success and failure for each rep. The four strategies were his current strategy (starting at 209 kg, then 217 kg, and finally 224 kg if he succeeded on the previous two), along with a more aggressive strategy, and two unique strategies we named the “hourglass” and “gold medal” strategies. After three iterations of the Markov chains, these were the probability distributions:

Weight lifted	Current Strategy	Aggressive Strategy	Hourglass Strategy	Gold medal Strategy
0 kg (failed all attempts)	0.000125	0.238	0.000125	0.512
209 kg	0.397	0	0.648	0
217 kg	0.531	0.578	0	0
224 kg	0.072	0.169	0.314	.384
230 kg	0	0.0152	0.038	0.096
230+ kg	0	0	0	0.008

In conclusion, we found that, in the 109 kg weight class, there is usually an advantage to being as heavy as possible. However, that advantage did not transfer over to Caine Wilkes particularly, although there may be other factors (such as age) that are specifically impacting performance at larger weights. Finally, especially if he competes in the 2024 olympics in Paris, Caine Wilkes may want to consider updating his clean and jerk strategy, depending on what his goals are for the competition.