

Portfolio: Advanced Stats

by Jan Sodoge

This portfolio presents an overview of techniques, features and methods to gain insights from existing, and often unused, data in European basketball. Before heading into the different statistical tools, let's clarify my perspective on the use and purpose of statistics in sport. Contrary to what the Moneyball mantra claims, statistics do not represent the holy grail to evaluating players or teams. Statistics, considered in a vacuum, can be horribly misleading. Consequently, better use them as a supporting tool for your eyes, an (additional) pair of glasses! There is not a single metric as detailed as what coaches see on the court themselves. Nevertheless, the power of data is to create knowledge of masses of game data about players, games and teams beyond where human brain capacity reaches natural limitations. Information provided by statistics and data is used best to check whether ones own hypotheses or to gain an initial evaluation using them.

Player Evaluation, Comparison & Development

Besides watching and scouting a player using your eyes, as a second perspective, advanced statistics can be approached to compare and evaluate players and check your personal hypotheses. Players can be evaluated, compared and analyzed in their personal development using metrics which each cover different aspects of their game. Using statistics as a second opinion here can be approached for roster planning, in-season player evaluation or scouting opposing players. Example give, insights on play-making abilities between two point-guards (see Figure 1) or shooting developments of shooting guards over the course of a season or their career can provide effective information. Similarly, the scribed approach can be transferred to teams as a unit.

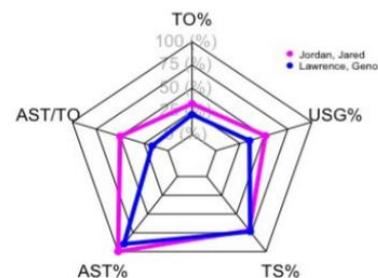


Figure 1: Radar chart to compare two players

Lineup and Player Synergies

Besides advances in statistical metrics for individuals, basketball remains a team sport. To treat it as a team sport using statistics, analyzing lineup and player combinations on court fosters assessing the collective level performances on court. Such a holistic approach helps to highlight synergetic effects between players on court: which players to best combine in the front court, who is beneficiary to a certain lineup, who is not? How does your team perform if #2 and #4 are on court and #1 and #3 benched? These questions can be best answered by considering lineup statistics. Using computational power, one is able to analyze all potential lineup combinations between 2,3,4 or five players. Thereby, hidden patterns, positive or negative lineups combinations can be identified. Given the huge amount of potential lineups, a task that a human mind can hardly take care of.

Spatial Shooting

Shooting and scoring with the basketball remains, besides all modern metrics, a core variable in traditional player evaluations. “This guy can shoot the basketball”. Currently available data, metrics and visualizations support more detailed information than commonly used approaches (FG%). These components offer important insight to shooting behavior, spatial patterns and efficiency. E.g., shot charts assist the recipient to analyze both volume and efficiency of a player from any given spot on the court (Figure 2).

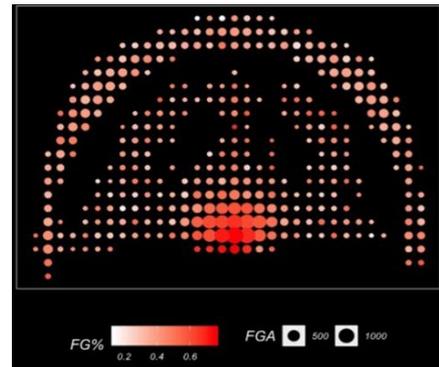


Figure 2: Shotchart displaying spatial patterns

In-game Evaluation

Yet presented features were shown to be used in an office or training context. But why not step analytics up the court, live, during competition? First experiments during the past season (run in cooperation with s.Oliver Würzburg) showed the potential to deliver intuitive data and visualizations to the bench via a personalized app on tablet or during halftime. In first experiments, information on lineup performance, fouling issues and player workload were used by the assistant coach to enrich and simplify the coaching.

Scouting database and data management

Scouting represents a major task, consuming substantial time within the coaching staff. For a productive and efficient scouting process throughout a longer timeframe, structured data storage and management are vital. Consequently, an individualized scouting database, including a user interface for scouting and displaying yet existing scouting data, guarantees multiple advantages. First, the database tool can be developed to fit individual scouting habits. This relates to the database structure i.e. What categories for a player/ team can be given scouting notes to, the user interface for writing scouting notes and displaying the database. Second, a web based tool can be used to allow you typing scouting notes from your smartphone while watching games or back in the office. Third, as the data is individually structured and stored in a single database, additional data sources e.g. Synergysports or open source data can be added. Thereby, the scouting database can integrate the other statistical driven approaches presented in this portfolio.

Roster Fit

Rather described to be in a state of research, this project tries to use methods from machine learning to predict a fit of players without these players never having played before on court together. Using such a method could potentially foster roster scouting and designing rosters or player additions during the season.