

Collegiate basketball tournaments, a complete systemic analysis

Systems Analysis and Design

Workshop N°1

Nicolás Romero Rodríguez - code: 20222020023

Xiomara Salome Arias Arias - code: 20222020028

Carlos Andres Celis Herrera - code: 20222020051

Universidad Distrital Francisco José de Caldas

Bogotá, Colombia

## Collegiate basketball tournaments, a complete systemic analysis

## Systems Analysis and Design

## Workshop N°1

**Competition Overview**

The competition chosen for this analysis revolves around predicting the results of this year's collegiate basketball tournaments of both women's and men's first divisions, for this purpose the participants are provided historical data of previous games of this league which will be used to predict the result of every possible team matchup of the teams participating in the tournament. Submissions for the contest are evaluated using the Brier score between the predicted probabilities and the actual game outcomes, additionally, every participant can participate with two different submissions of predictions.

The Brier score is a metric used to measure the accuracy of probabilistic predictions. This score is calculated as the mean squared difference between the predicted probability and the actual outcome. It is defined as:

$$BS = \frac{1}{N} \sum_{t=1}^N (f_t - O_t)^2 \quad (1)$$

Where N is the number of predictions,  $f_t$  is the probability of the event occurring, and it corresponds to the actual outcome, it's set to 1 if the event happened or 0 if it didn't. The closer the brier score value is to 0, the more accurate the prediction was. The datasets given to make the prediction systems are given in csv format, both genders have separate datasets and contain extensive information on the teams records since 1984 separated in five sections.

- Section 1: Contains basic teams' information including, Team ID's and team names, tournament seeds since 1984 -85 season, final scores of 3 different tournaments since 1984 and season level details including dates and regions.
- Section 2: Contains game-by-game stats at team level, including data such as free throws attempted, assists, blocks, three pointers or rebounds made by each team

across 3 different tournaments since 2003 for men's division and 2010 for women's division.

- Section 3: Contains a list of cities where games have been played for the three tournaments.
- Section 4: Contains weekly team rankings (only available for men's teams) from various top rating systems since the 2003 season.
- Section 5: Contains supporting information including coaches, conference affiliations, bracket structures, and game results of additional tournaments

Despite the above information there are some additional constraints to the data presented, as there are differences in the amount of data that exist for the women and men's divisions, aside from that some of the data provided is incomplete as earlier women's data is not obtainable in these datasets, for example an approximate of 1.5% of games data from the 2010 to 2012 is unavailable in sections 2 and 3. ?

### System Analysis report

Figure 1 shows all elements and interconnections of a basketball tournament, the following report will explain all of these components.

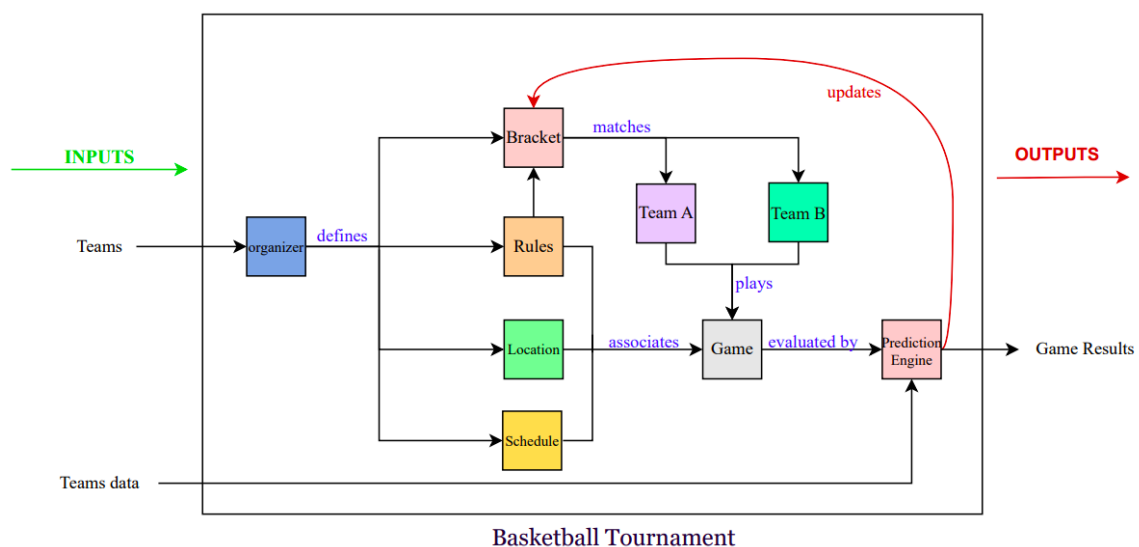


Figure 1. Collegiate basketball tournament represented as a system.

## Inputs

In this system the inputs are represented by the teams that are selected to play in the tournament, in this specific case there are 64 teams in each category. Also, teams data contains all relevant teams information (statistics, history, performance, etc).

## Outputs

It represents the specific match in which two teams, Team A and Team B, face each other as part of the basketball tournament. This game is organized based on the information provided by the Schedule and Location, defined by the organizer. Once the match is played, the result is used to determine the Match Probabilities, which in turn updates the system and helps to generate the tournament results.

## Elements

- **Organizer:** In this context, the organizer is who defines the tournament organization, as brackets, rules, location and schedule of the games.
- **Bracket:** Table or diagram that organizes the matchups in a tournament. It allows predicting which teams will advance to the next round.
- **Rules:** Rules governing the operation of the tournament, from the way in which players must play in each match, and the way in which teams will be ranked or eliminated.
- **Location:** Place where the teams face each other, the city and the field are considered as these.
- **Schedule:** Time and date of the games.
- **Team A and Team B:** Organized group of people working together to beat the opposing team.
- **Game:** Represents the specific match in which two teams, Team A and Team B, face each other as part of the basketball tournament. This match is organized

based on the information provided by the Schedule, Rules and Location, defined by the organizer. In essence, this element is all the information about the game.

- **Prediction engine:** The Prediction Engine is the system component that is responsible for estimating the probability of victory between two teams in a simulated match. To make this prediction, it receives as input an object that represents the match (Game), together with the corresponding data of each team. Through a predictive model, the engine analyzes the relevant information (such as offensive and defensive efficiency, ranking, etc.) and generates a numerical probability that the first team (Team A) will win the match.

## Relationships

- Organizers define the rules, bracket, location, and schedule of the games based on the participating teams.
- The bracket matches teams in pairs according to the organization's setup.
- Rules, location, bracket, and schedule are associated with a game.
- Team A and Team B play in the game.
- The game determines the results.
- The Prediction Engine updates the bracket to determine future matches in the tournament.
- The game containing all the information is evaluated by the prediction engine.

## Processes

The tournament starts with 4 matches that decide the last 4 teams to classify to the official tournament, completing a total of 64 participant teams; this part of the tournament is also known as "The first four" (Wilco (2025)) Once the teams that will participate in the tournament are selected, each team is assigned to a bracket that decides which team they will play against initially. These brackets are not selected

randomly, as they are carefully selected using a seeding system that uses the following logic.

Teams are seeded from 1 to 16 within each of the four regional brackets of the US (East, West, South, Midwest). The Selection Committee determines seeds based on factors such as a win-loss record, the strength of the schedule, quality wins, and overall performance throughout the season.

Wilco (2025) Once the seeds are decided, the brackets are set up so that the highest-seeded teams face the lowest-seeded teams in the first round, for example:

- 1 vs. 16
- 2 vs. 15
- 3 vs. 14
- 4 vs. 13
- 5 vs. 12
- 6 vs. 11
- 7 vs. 10
- 8 vs. 9

This structure rewards higher-seeded teams by giving them theoretically easier opponents in the first round. After the first round, winners move on to the next round, and matchups continue according to the bracket structure, maintaining the principle that the highest remaining seed faces the lowest remaining seed as much as possible, the purpose of this system is to reward teams that had a better performance during the season while still allowing lower-seeded teams a chance to compete.

All of the tournament's matches are decided by a set of rules decided by the NCAA and are as follows: Games for the men's division consist of two 20 minute halves, while games for the women's division consist of four 10 minute halves. If at the end of this time the game is tied a 5-minute overtime period is played, this overtime

repeats until a winner is determined. *March Madness rules: Why the men play halves and the women play quarters* (2025) Each team gets 4 timeouts per game (3 one-minute timeouts and 1 30-second timeout) with an additional full timeout for every overtime period played.

The tournament has a single-elimination format, which means that once a team loses a game the team is disqualified. The format for each round is arranged similarly to a binary tree, the First Round is played by 64 teams, the Second Round consists of 32 teams, the third round also known as the Sweet 16 is played by 16 teams, then comes the eighth round also known as the Elite Eight, after this comes the Final Four, and finally, the Championship Game, which determines the national champion.

Scoring in March Madness follows typical college basketball rules. A field goal made inside the three-point line counts as two points, while a successful shot from beyond the arc is worth three points. Free throws are worth one point each. Teams accumulate fouls throughout each half, and after committing seven fouls, the opposing team is awarded a one-and-one free throw (the player must make the first to earn a second shot). When a team reaches ten fouls, the opponent automatically gets two free throws, known as the double bonus. If a player commits five personal fouls, they are disqualified from the game. Technical fouls result in two free throws and possession for the opposing team. The shot clock for both men's and women's games is set at 30 seconds. NCAA.com (2025)

Locations for matches are decided years in advance by the NCAA, games for these tournaments are usually hosted in large venues in order to accommodate the largest number of people possible, these selected venues are spread across the whole country to prevent the teams from having to travel long distances for any match. The Final Four and final game are usually held in major cities with large arenas or domes to accommodate a large audience. NCAA.com (2025) Once the bracket is set, the NCAA schedules games in a way that maximizes viewership, for this purpose, games are scattered throughout the day, so that the audiences can watch multiple games in a single day. The tournament is played over three weekends, with two rounds occurring

each weekend, this spacing aims to give the teams enough time to travel to their respective locations for the next match and allow the players to have enough rest so that they can play in optimal form.

### **Complexity and sensitivity**

The teams themselves are their own subsystem, which increases the complexity of the whole system, and causes that even things like the performance of a single player may affect the overall result of the tournament.

Social and political circumstances may also affect the system if, for example, an ongoing protest prevents a team from arriving on time to a game then the game may have to be delayed, if the delay is a couple hours long (as seen in other important sport events in the USA).

External factors like weather and temperature may impact some of the players performance, or even if the delay makes it so that a team has to warm-up more than once previous to the game. So, in general anything that may cause a logistical problem that affects the expected development of the game is a limitation of the tournament.

Another factor to consider is the way that the playstyles of each team collide with their rivals, as even if a team's strategy is organized and solid, their rivals may have a strategy that counters their strategy, because of this, teams should be adaptable in order to prevent such situations.

In the system, the number of teams that enter is important because if any of them do not show up, the organizer will have to change the tournament format. This would affect the bracket, location, rules and schedule, since the organizer would have to make changes to the matchups, adjust dates and times of the games. In addition the games would no longer be the total expected by the audience causing negative effects outside the environment. Finally, the results could not take into account the statistics of the teams that do not attend the tournament. This change in the entries and the repercussions it would have on the level of each element demonstrates the high sensitivity of the system.



On the other hand, there could be sensitivity to changes in the environment. These changes include: climate, social, economic and institutional changes. Each of these changes can affect the organizer and as in the changes presented by the inputs happen again in this case. Moreover, it can be considered that an element with a high sensitivity is the organizer because a change in its entry or in the elements directly related to it leads to a readjustment in the format of the tournament.

### **Chaos and randomness**

There are many factors that can influence the outcome of a tournament match and that can occur unexpectedly, so they could influence the final outcome of the tournament. These chaotic factors could affect game predictions, since they are unmeasured and unexpected in the middle of a real-life game:

- **Chaotic Feedback in the Bracket:** The prediction engine relies on structured data that does not account for real-world chaotic factors such as injuries, referee mistakes, or emotional pressure. Because these unpredictable elements are not included, the simulated system cannot replicate real-world chaos. However, this limitation is what makes real tournaments inherently chaotic and unpredictable. In contrast, the simulation, though built on deterministic rules, becomes unpredictable in its own way, due to how each prediction influences the next. The bracket evolves dynamically, showing how complexity emerges even without modeling real-world randomness.
- **Performance influenced by the public:** The public is an element of the system's environment that can be influential on the team, this unexpected interaction can change the players' mood, generate additional pressure or boost performance in a surprising way.
- **Personal circumstances in players' personal lives:** Personal circumstances such as familiar difficulties, personal worries, grudges among team members or even a bad night of may interfere in the team's preparations and can impact the

players' performance in the match, and even if only one of the players' is affected by a personal problem the team's general performance can be heavily impacted and this factor could make the difference between winning and losing a game.

- **Court illumination:** Despite being played in enclosed gyms and as such not specially affected by rain or intense sunlight, the gym's illumination may vary from court to court, it also may be different if the lights were recently changed, this may impact a game's result in the sense that illumination may give the players more or less visibility of their surroundings affecting their precision on passes or throws or could momentarily flash a player not used to a particular bright of light.
- **Injuries and fatigue:** A player getting injured during a game forcing a change, or even a player pushing through an injury could affect a team's performance specially in a fast-paced, physical game such as basketball.
- **Referee bias and mistakes:** A referee missing or making incorrect calls may impact the flow of the game or give advantage to a particular team, aside from impacting the teams' morale potentially causing frustration on players that may eventually lead to confrontations which of course may greatly affect a match's result.

## Conclusion

A basketball tournament is composed by many elements and subsystems, very reliant on logistics for it to work correctly, and with every element being capable of affecting the results of the matches in a way or another.

Tournament scheduling may be affected by changes in entries and environment. Since the organizer is very sensitive it leads to the other elements being affected by this as well.

Tournaments are dynamic ecosystems, with teams and players constantly adapting to large range of possibilities in order to win the tournament.

One of the strengths of the system is that its elements have synergy to a large

extent. In a concrete measure we have that the sum of the organizer with the rules leads to play a clean match without chaos. Now in a much broader vision, we have to define who is the winner of a match or tournament, this result is achieved by the contribution of each of the elements of the system.

The result obtained by the game and compared with the prediction obtained by the prediction engine, may have more than one interpretability if these two are different. It is difficult to determine whether the prediction was well founded or the model may have been lucky.

## References

*March madness rules: Why the men play halves and the women play quarters.* (2025).

Retrieved from

<https://www.espn.com/mens-college-basketball/story/id/44344981/march-madness-rules-why-men-play-halves-women-do-quarters>

(ESPN.com)

NCAA.com. (2025, March 31). *2025 march madness: Men's ncaa tournament schedule, dates.* Retrieved from

<https://www.ncaa.com/news/basketball-men/article/2025-03-30/2025-march-madness-mens-ncaa-tournament-schedule-dates> (NCAA.com)

Wilco, D. (2025, January 22). *What is march madness: The ncaa tournament explained.* Retrieved from

<https://www.ncaa.com/news/basketball-men/bracketiq/2025-01-22/what-march-madness-ncaa-tournament-explained> (NCAA.com)