Predicting the 2021 Futsal League

Are you excited to find out the results of this winters futsal league? An exciting match this past weekend saw LaMancha fight off dark horses Kaghani to remain the only undefeated Joy team left. Both sides have guaranteed themselves a playoff spot already, but do they have what it takes to win the championship? I did the math to find out, so that you don't have to wait any longer!

Playoff Results-1-1

	1	2	3	4
Sala Boys Black	63.68%	24.02%	10.08%	2.22%
JOTP LaMancha	19.01%	38.59%	33.16%	9.24%
JOTP Kaghani	15.88%	29.49%	42.12%	12.51%
KFA Boys	0.65%	2.84%	5.08%	16.63%
Blackhawks 2006 Black	0.46%	2.91%	5.16%	30.89%
FC International	0.31%	2.03%	4.06%	24.76%
Joga Bonito Rio	0.01%	0.11%	0.34%	3.73%
JOTP Kiko	0.00%	0.01%	0.00%	0.01%
Sala Boys Blue	0.00%	0.00%	0.00%	0.01%

This table shows the probabilities that a team will finish the playoff champion, runner up, third place, or fourth place

2021 Regular Season Results

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Sala Boys Black	78.61%	20.87%	0.52%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
JOTP LaMancha	21.37%	46.80%	31.83%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
JOTP Kaghani	0.02%	32.33%	67.65%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Blackhawks 2006 Black	0%	0%	0%	25.20%	21.83%	32.75%	19.31%	0.82%	0.07%	0.02%	0%	0%	0%	0%	0%	0%	0%
KFA Boys	0%	0%	0%	39.42%	21.33%	20.89%	6.14%	0.26%	4.57%	7.39%	0%	0%	0%	0%	0%	0%	0%
FC International	0%	0%	0%	31.16%	42.89%	18.56%	3.61%	2.51%	1.04%	0.23%	0%	0%	0%	0%	0%	0%	0%
Joga Bonito Rio	0%	0%	0%	4.19%	12.88%	15.12%	14.59%	25.43%	21.92%	5.87%	0%	0%	0%	0%	0%	0%	0%
Sala Boys Blue	0%	0%	0%	0.01%	0.63%	7.90%	36.11%	14.05%	15.05%	26.25%	0%	0%	0%	0%	0%	0%	0%
JOTP Kiko	0%	0%	0%	0.02%	0.39%	4.37%	17.55%	23.19%	12.21%	42.27%	0%	0%	0%	0%	0%	0%	0%
JOTP Garganica	0%	0%	0%	0%	0.05%	0.41%	2.69%	33.74%	45.14%	17.95%	0.02%	0%	0%	0%	0%	0%	0%
St Paul Blackhawks 04 Red	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.02%	34.27%	65.70%	0.01%	0%	0%	0%	0%
Prior Lake Typhoon	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	65.71%	31.30%	2.98%	0.01%	0%	0%	0%
Harambee HS Boys 1	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3.00%	55.11%	6.51%	28.54%	6.74%	0.10%
JOTP Verata	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	40.18%	13.67%	35.85%	9.98%	0.32%
JOTP Murciana	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.39%	70.81%	26.22%	2.55%	0.03%
Hawks	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1.33%	7.04%	8.74%	80.65%	2.24%
JOTP Girgentana	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1.96%	0.65%	0.08%	97.31%

In this table you can see the probability that your team will finish in 1st-17th!

The Process of Predicting

I'm sure you are wondering how I got to these results. The short answer is I used a computer to simulate all remaining matches (including the playoffs) 10,000 times! Then I can see for example how many times out of 10,000 LaMancha finished in 1st place to find the probability that they will finish in 1st place in real life!

But how do you simulate a match in the first place? Let me take you through the process.

Predicting the Tennessee Feinting / Sala Girls Match

The girls have a big match this weekend against Sala Girls, one that will decide the winner of the league! To see if we can predict the outcome of this match we will need a few things..

1) The "Attack Score" and "Defense Score" of each team. How good are both of these teams at attacking and defending compared to the average team? We calculate this:
   Attack Score = Goals per Game / League Avg Goals per Game
   Defense Score = Goals Allowed per Game / League Avg Goals Allowed per Game

	Attack Score	Defense Score
Tennessee Feinting	1.58	0.54
Sala Girls	1.44	0.34

2) With these scores we can calculate the number of goals each team should expect to score against each other (we'll call it Expected Goals)
   Team 1 Expected Goals = Team 1 Att Score * Team 2 Def Score * Avg Goals/Game
   Team 2 Expected Goals = Team 2 Att Score * Team 1 Def Score * Avg Goals/Game

	Expected Goals
Tennessee Feinting	2.98
Sala Girls	4.33

3) **ADVANCED MATH ALERT** We use something called the Poisson distribution to calculate all possible scoreline probabilities

Sala Girls Score

		0	1	2	3	4	5	6	7	8	9	10	11
	0	0.1%	0.3%	0.6%	0.9%	1.0%	0.8%	0.6%	0.4%	0.2%	0.1%	0.0%	0.0%
	1	0.2%	0.9%	1.9%	2.7%	2.9%	2.5%	1.8%	1.1%	0.6%	0.3%	0.1%	0.1%
	2	0.3%	1.3%	2.8%	4.0%	4.3%	3.8%	2.7%	1.7%	0.9%	0.4%	0.2%	0.1%
Tennessee	3	0.3%	1.3%	2.8%	4.0%	4.3%	3.7%	2.7%	1.7%	0.9%	0.4%	0.2%	0.1%
Feinting	4	0.2%	1.0%	2.1%	3.0%	3.2%	2.8%	2.0%	1.2%	0.7%	0.3%	0.1%	0.1%
Score	5	0.1%	0.6%	1.2%	1.8%	1.9%	1.7%	1.2%	0.7%	0.4%	0.2%	0.1%	0.0%
	6	0.1%	0.3%	0.6%	0.9%	1.0%	0.8%	0.6%	0.4%	0.2%	0.1%	0.0%	0.0%
	7	0.0%	0.1%	0.3%	0.4%	0.4%	0.4%	0.3%	0.2%	0.1%	0.0%	0.0%	0.0%
	8	0.0%	0.0%	0.1%	0.1%	0.2%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%
	9	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	10	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

The most probable score lines include 3-3, 4-3, 4-2 implying another very close game. The last meeting ended 5-4 to the Tennessee Feinting. If we add up the probabilities we get a 24.4% chance of a JOTP win, a 13.4% chance of a draw and 62.2% of a Sala win. Sala's defensive record of only allowing 17 goals in 9 games make them mathematical favorites in this matchup, but we will have to tune in this weekend to find out if JOTP can repeat their result against Sala or if they tense up and fall over on the court, just like the goat they represent!

Predicting the Entire League

Now that we know how to predict a match, all we need to do is do that for all remaining fixtures to get the final league table. In the example below LaMancha was able to overcome Sala to win the regular season, and went on to beat them again in the finals. Kaghani defeated 06 Blackhawks for third place. But this is just one iteration! Now we do this 9,999 more times to figure out just how likely any given result is!
If you play for Kaghani, don't get down over your first loss - the math says you still have a 45.37% chance of reaching the finals! And if you play for Kiko don't give up hope, you are still in playoff contention! You guys made the playoffs in 2 out of the 10,000 simulations, and even had 1 finals appearance!!

Team	GF	GA	GD	Pts
JOTP LaMancha	97	50	47	30
Sala Boys Black	82	27	55	27
JOTP Kaghani	90	40	50	27
Blackhawks 2006 Black	76	57	19	21
FC International	59	57	2	19
KFA Boys	80	58	22	18
Joga Bonito Rio	71	57	14	18
Sala Boys Blue	62	48	14	18
JOTP Garganica	73	69	4	18
JOTP Kiko	62	49	13	16
Prior Lake Typhoon	56	60	-4	12
St Paul Blackhawks 04 Red	60	67	-7	12
JOTP Verata	39	77	-38	6
JOTP Murciana	51	111	-60	4
Harambee HS Boys 1	37	71	-34	3
Hawks	36	82	-46	3
JOTP Girgentana	34	85	-51	1

An example of one league simulation. One of the 10,000 we use to find probabilities!

Limitations

Its important to know your limitations when calculating things like this, because the last thing you want is to be confident in your work and then end up completely wrong! So here are a list of possible issues with my methods and what I did to try to make things as accurate as possible.

1)   Fair Schedules - In the boys league, not everyone's schedules are "fair" since theres 17 teams and only 10 games. For example Sala Boys Black have had the second easiest schedule since their opponents so far average 1.35 points per game. Meanwhile LaMancha, the only other undefeated team have a had a bit harder of a schedule - their opponents average 1.47 points per game. And poor 04 Blackhawks' opponents have averaged a whopping 1.88 points per game. I was

able to take this into consideration by incorporating it into the "Attack Score" and "Defense Score" calculations.

2) Sportsmanship - Of course everyone knows Joy has the best sportsmanship, and that Joy teams are less likely to be running up the score when its unnecessary. But the math favors teams that run up the score, so how do you combat this? I used the leagues restriction where you cannot win by more than 7 goals (a 10-1 win turns into an 8-1 win, etc). This is an ok fix, but still leaves some questions in the accuracy.

3) Team Set Restrictions - If a Joy team knows they will win the game, or is leading by enough, the team will put a restriction on scoring (cycle before you score, skill moves before you score, etc). It is possible but not guaranteed other teams will do this as well, meaning the math will again favor "try hard" teams. This is a limitation I don't currently have the tools or the data to solve.

It's impossible for the the math to be completely perfect - we can only strive to get as close to reality as possible. This model does a good job of that but it is very important to always remember it is not 100% accurate.

For those of you who really want to go above and beyond I'll try and link the Python code I used to do the simulations. The code I wrote myself, but the methods of prediction are common ideas, here is a link for their application in the Premier League. https://medium.com/@DanielAKeane/soccer-statistics-using-poissons-distribution-to-predict-results-of-upcoming-games-a-real-337f383889c9

Good luck and happy coding!