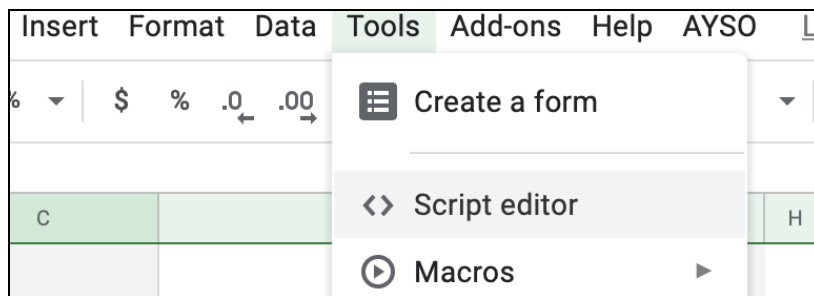


Gradient-descent least-squares optimization to allocate players to soccer teams

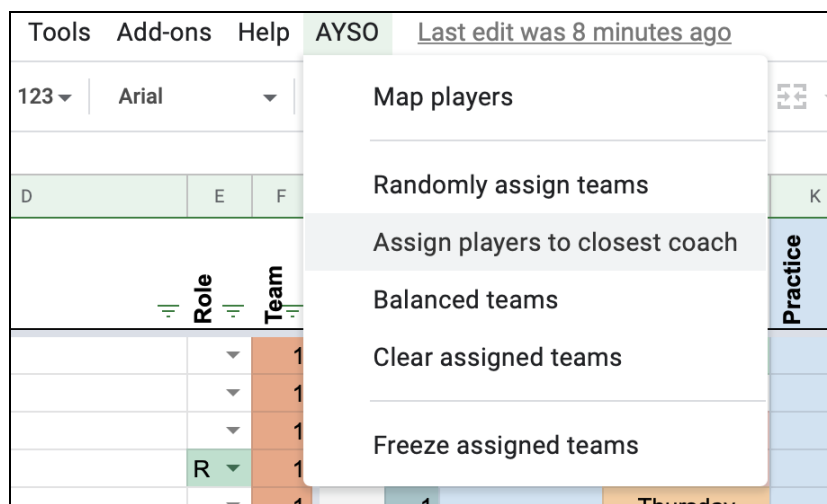
Shanti Rao, Aug 5, 2020

Each division coach administrator (DCA) gets a [spreadsheet](#) with information on the players and volunteers in her division (this one has been anonymized and the names scrambled). The gray columns have data that comes from the registration form. The blue columns are provided from other algorithms. The spreadsheet also has address, school, and contact info. In younger divisions, the DCA can use that information to assign players for teams.

For 6U, the players are allocated to the coaches that live closest to them. To see it work, go to the Tools menu and open the Script Editor. It'll show you some scary warnings. Just click through and accept everything.



When you see the script editor, run the “OnOpen” script and go back to the spreadsheet, where you’ll see a new AYSO menu. This will happen automatically the next time you open the spreadsheet. It won’t work for you here because I’ve cleared the Address and City. But feel free to edit the geographic boundaries in the “Code.gs” file and make up some addresses to try it. Otherwise, keep reading.



Parents are asked how they want to volunteer for their child's team. Some of them answer "Coach" or "Referee". The DCA will call the parent, confirm, and then mark the Role column appropriately. Sometimes, a family will do more than one role for a team. In this case, the DCA types in something like "AC,R" for Assistant Coach and Referee. Google doesn't know how to validate it, but the software can handle it. "Q" is a special category of parents who shouldn't be on the same team, and the algorithm will automatically distribute them throughout the division. Not everyone follows through on the volunteer role that they write on the signup form, so it's important for a DCA to get to know the volunteers and make the decision of what they're going to do.

You could, hypothetically, ask is what day the player would prefer not to practice. This doesn't go into the algorithm yet, but the DCA can use it to manually assign players to teams. To assign a player to a team, the DCA simply types a number into the Team column. Update: It turns out not to be very useful, because what people say in May often differs from what they want in September.

Some players are assigned to team -1, indicating that they withdrew and should be ignored.

Pool is used to separate a division into different flights, so that the teams within each flight would be balanced. This feature is not yet implemented.

Practice is the day a coach would like for practice. In a perfect world, we would use this to avoid conflicts. This feature is not yet implemented.

Siblings is a comma-delimited list of the AYSO IDs of other players who need to be on the same team.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	AYSO ID	Form	Player First Name	DCA notes	Role	Team	Status	Pool	Siblings	Avoid	Practice	Conflicts	Volunteer
13	124		Andrea			1		1		Monday			I cannot or would prefer
14	125		Annabelle			1		1		No			I cannot or would prefer
15	126		Annalina			1		1		Friday			Referee
16	127		Annelise		R	1		1		Friday			No
17	128		april			1		1		Thursday			Manager
18	129		Ava		M	1		1		No			I cannot or would prefer
19	130		Ava			2		1		Wednesday			Help with tournaments a
20	131		AVA			2		1		Friday			I cannot or would prefer
21	132		Beatrice			2		1		Friday			I cannot or would prefer
22	133		Belisiah			2		1		Friday			Help with tournaments a
23	134		Bella A		HC	2		1		Wednesday	Thursday		Coach
24	135		Bella B			2		1		Friday			Staff the tent on Saturda
25	136		Brianna			2		1		Friday			No
26	137		Camryn			2		1		Wednesday			No

DCAs are also provided with assessment data. Goals Scored and Quarters in Goal come from the previous season, for returning players for whom we have statistics. [Rating and Strengths and Weaknesses are provided by the previous season's coach](#), but the rating is adjusted by the

coach's effectiveness in the previous season. Players on losing teams who were highly rated have their score reduced, and players on winning teams who were underrated have their score increased. Don't interpret the rating as an assessment of the player's worth -- an undervalued player will be placed on a team with better players, leading to a better record in the next season. (At the end of a season, a team's average player rating should be roughly proportional to the team's win/loss record. To the extent that it's not, the coach's assessment is handicapped. We've explored many other post-season metrics, like weighted averages of goals for and against, or the coach's assessment of their own child, and found nothing that works better than 3 points for a win and 1 point for a draw.)

Ability is a parent's assessment of their child. It says more about the parent than it does about the child, which can be useful information for coaches.

History	Rating	OLD Rating	Overall	Goals Scored	Quarters In Goal	All Star	Strengths and Weaknesses	Ability	Birth Month	Birth Year
Yes, in another AYSO region								Good at soccer	9	2007
Yes, in Region 13	8.0	8				X		AYSO All-Star player	8	2007
Yes, in Region 13	6.7	7		5		A	She is young for division	EXTRA or Club player	5	2007
Yes, in Region 13	2.4	3						Plays occasionally with friends	10	2006
Yes, in Region 13	6.2	7		8		A		EXTRA or Club player	9	2006
Yes, in Region 13	2.5	3					Insists to play her way	Enthusiastic about soccer	5	2007
Yes, in Region 13	7.9	10						Good at soccer	3	2006

Birth Month and Birth Year are to balance teams and avoid the "All Hockey Players are Born in January" effect. Prior to balancing teams by birth month, we found that player rating, and retention, was skewed significantly to the oldest two birth months in the cohort. Implementing this distribution roughly coincided with shifting the age groups. You can still see a rough correlation on the Stats page between age and a coach's rating, but at least there are plenty of November and December kids. The January kids are all playing Club or on the EXTRA team. June and July are depleted in this particular cohort from the legacy of the July/August age cutoff. Rating aside, we're doing much better at retaining the November + December kids now than we used to for July + August.

The "All Star" category is to flag players from the previous season's travel teams, so they can be distributed throughout the division.

How to use it

Flag as many parents as you can as HC (Head Coach), AC (Assistant Coach), R (Referee), or M (Manager). There will be as many HCs as there will be teams. The other categories will be uniformly distributed according to the preferences you set up on the Balance tab.

If you want to assign some players to a specific team, you can do that by putting a positive number in the Team column. Those players will stay put and everyone else will move around. You can use this to distribute some players, then freeze those assignments, and then distribute a new batch of players.

About the balance tab ...

The algorithm seeks to minimize a cost function. In this case, we have collected 18 heuristic rules that we think should be roughly the same for each team.

Importance is how to weight that particular heuristic in the cost function

Field is which column is used to calculate the heuristic.

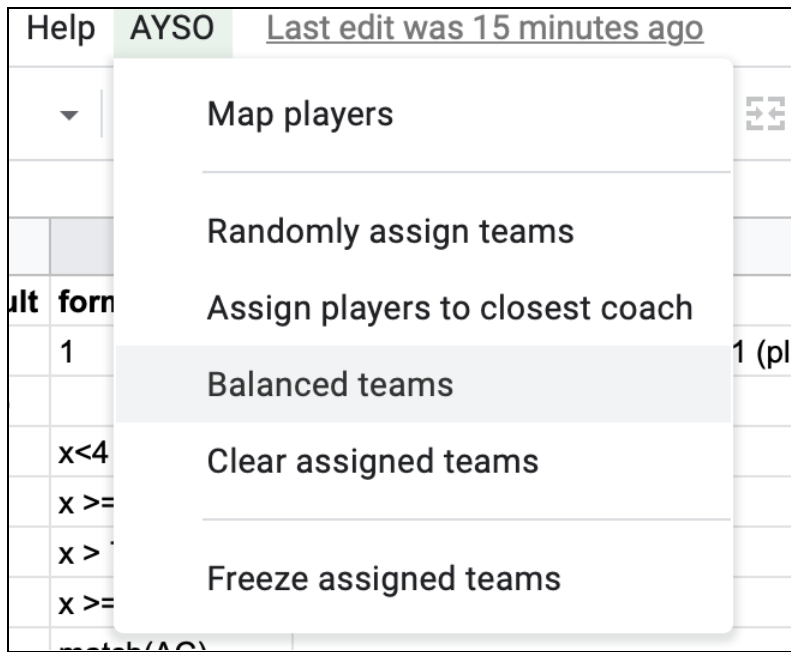
Combine is whether it's the number of players in that category that should be matched, or the average of a numeric score that should be matched.

Limit is whether there is a hard upper limit on the number of players in that category. In this example, no team can have more than $\text{ceiling}(\text{players}/\text{teams})$ players. Superstar players and ACs are similarly allocated, so that no parent complains, "that other team has 2 more great players than ours!"

1	name	importance	field	combine	limit	default	formula
2	players	1	AYSO ID	count	yes		1
3	rating	1	Rating	average		5.5	
4	below	2	Rating	count			$x < 4$
5	average	0.5	Rating	count			$x \geq 4 \ \&\& \ x \leq 7$
6	above	2	Rating	count			$x > 7$
7	superstar	1	Rating	count	yes		$x \geq 8$
8	ac	1	Role	count	yes		match(AC)
9	ref	1	Role	count			match(R)
10	manager	0.5	Role	count			match(M)
11	useful	1	Role	count			match(R,AC)
12	early	1	Birth Month	count			$x < 5$
13	late	1	Birth Month	count			$x > 8$
14	bestallstar	4	All Star	count			match(X,S)
15	goodallstar	1	All Star	count			match(R,A)
16	older	2	Birth Year	count			$x < 2006.5$

To run the algorithm, go to the AYSO menu and choose Balanced Teams. It will find a solution, and then keep running to try to find a better solution. Google Apps Scripts limits execution time to 5 minutes, so the code is written to resume where it left off if you do it again and again. Just wait for it to finish before you start it again!

When you're satisfied with the result (look at the box scores on the Teams tab), run "Freeze assigned teams" and those players will stop moving around.



Feel free to use and adapt the source code, and pay it forward!