

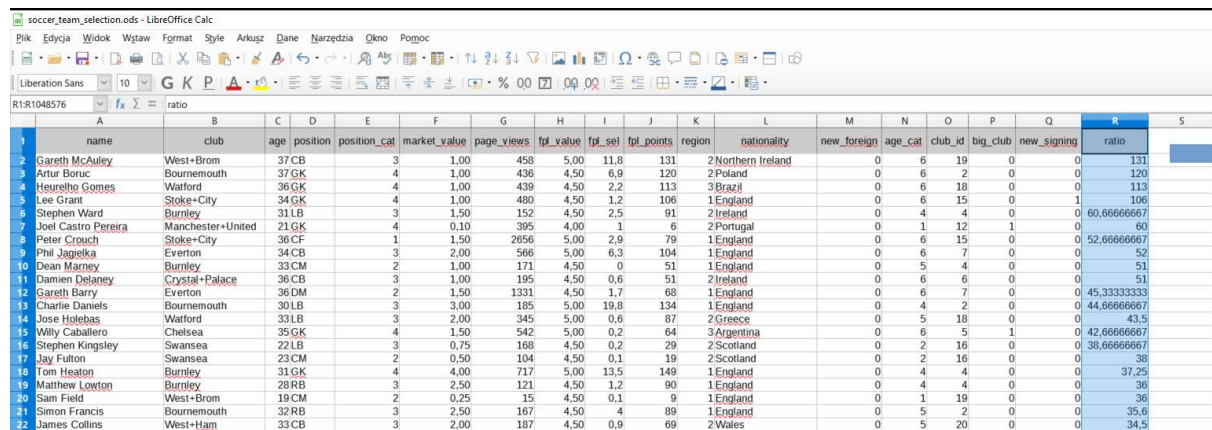
Football team optimization with Solver

Aim of this task was to complete a team of soccer players, maximizing teams total score while staying within given budget - 100 MLN. As a dataset I choose English Premier League Players Dataset from 2017/18 from kaggle <https://www.kaggle.com/mauryashubham/english-premier-league-players-dataset>.

Crucial features for performing optimization were:

- market_value,
- position_cat, 1 - attacker, 2 - midfielder, 3 - defender, 4 - goalkeeper (there were made little simplification - I didn't distinguish left/right defenders and midfielders etc. to avoid huge 0/1 table, let's assume that players are flexible enough to play on both sides of the pitch),
- fpl_points - Fantasy Premier League Points - points accumulated over the previous season.


First step was to calculate players ratio showing points per 1 MLN for each player from the dataset by performing simple division $fpl_points / market_value$, column S on screenshot below:



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
	name	club	age	position	position_cat	market_value	page_views	fpl_value	fpl_sel	fpl_points	region	nationality	new_foreign	age_cat	club_id	big_club	new_signing	ratio	
1	Gareth McAuley	West+Brom	37	CB	3	1.00	458	5.00	11.8	131	2	Northern Ireland	0	6	19	0	0	131	
2	Artur Boruc	Bournemouth	37	GK	4	1.00	436	4.50	6.9	120	2	Poland	0	6	2	0	0	120	
3	Heurelho Gomes	Watford	36	GK	4	1.00	439	4.50	2.2	113	3	Brazil	0	6	18	0	0	113	
4	Lee Grant	Stoke+City	34	GK	4	1.00	480	4.50	1.2	106	1	England	0	6	15	0	1	106	
5	Stephen Ward	Burnley	31	LB	3	1.50	152	4.50	2.5	91	2	Ireland	0	4	4	0	0	60.66666667	
6	Joel Castro Pereira	Manchester+United	21	GK	4	0.10	395	4.00	1	6	2	Portugal	0	1	12	1	0	60	
7	Peter Crouch	Stoke+City	36	CF	1	1.50	2656	5.00	2.9	79	1	England	0	6	15	0	0	52.06666667	
8	Phil Jagielka	Everton	34	CB	3	2.00	566	5.00	6.3	104	1	England	0	6	7	0	0	52	
9	Dean Marney	Burnley	33	CM	2	1.00	171	4.50	0	51	1	England	0	5	4	0	0	51	
10	Damien Delaney	Crystal+Palace	36	CB	3	1.00	195	4.50	0.6	51	2	Ireland	0	6	6	0	0	51	
11	Gareth Barry	Everton	36	DM	2	1.50	1331	4.50	1.7	68	1	England	0	6	7	0	0	45.33333333	
12	Charlie Daniels	Bournemouth	30	LB	3	3.00	185	5.00	19.8	134	1	England	0	4	2	0	0	44.66666667	
13	Jose Holebas	Watford	33	LB	3	2.00	345	5.00	0.6	87	2	Greece	0	5	18	0	0	43.5	
14	Willy Caballero	Chelsea	35	GK	4	1.50	542	5.00	0.2	64	3	Argentina	0	6	5	1	0	42.66666667	
15	Stephen Kingsley	Swansea	22	LB	3	0.75	168	4.50	0.2	29	2	Scotland	0	2	16	0	0	38.66666667	
16	Jay Fulton	Swansea	23	CM	2	0.50	104	4.50	0.1	19	2	Scotland	0	2	16	0	0	38	
17	Tom Heaton	Burnley	31	GK	4	4.00	717	5.00	13.5	149	1	England	0	4	4	0	0	37.25	
18	Matthew Lowton	Burnley	28	RB	3	2.50	121	4.50	1.2	90	1	England	0	4	4	0	0	36	
19	Sam Field	West+Brom	19	CM	2	0.25	15	4.50	0.1	9	1	England	0	1	19	0	0	36	
20	Simon Francis	Bournemouth	32	RB	3	2.50	167	4.50	4	89	1	England	0	5	2	0	0	35.6	
21	James Collins	West+Ham	33	CB	3	2.00	187	4.50	0.9	69	2	Wales	0	5	20	0	0	34.5	

Then all 460 players were sorted descendingly by this ratio value. (Rumors told that Solver is able to manipulate up to 200 cells but later it came out that this one in LibreOffice does well even for 460 so sorting wasn't that necessary).

Later, assignments table was created to allow solver to manipulate with binary values which means 0 - player is not selected, 1 - player is selected. Following columns named accordingly to position abbreviations. They were fulfilled with 1 if player from specific row played on that position and 0 otherwise.

T	U	V	W	X	Y	Z
	<u>assignment</u>	<u>AT</u>	<u>MID</u>	<u>DEF</u>	<u>GK</u>	
	1			1		
	0				0	
	0				0	
	0				0	
	0			0		
	0				0	
	0	0				
	0			0		
	0		0			
	0			0		
	0		0			
	1			1		
	0			0		
	0				0	
	0			0		
	0		0			
	1				1	

All columns from U to Y were summed to allow Solver dynamically check if it meets the conditions. Moreover budget upper bound was given to hold total expense under 100 MLN. All constraints are listed in the table below:

Z	AA	AB	AC	AD	AE
	<u>features</u>	<u>constrains</u>	<u>relation</u>	<u>result</u>	
	<u>points</u>		<u>maximize</u>	1581	
	<u>expense</u>	100	>=	98	
	<u>players</u>	11	=	11	
	<u>attackers</u>	2	=	2	
	<u>midfields</u>	4	=	4	
	<u>defenders</u>	4	=	4	
	<u>goalkeeper</u>	1	=	1	

Finally constraints were given to Solver tool. Assignment cells U2-U462 were limited to binary values only, rest values were limited to total number of players in team and at specific positions. Of course optimization was performed in terms of maximizing total score of whole team.

Solver

Komórka docelowa:

Zoptymalizuj wynik pod kątem: ☒ Maksimum ☐ Minimum ☐ Wartość

Poprzez zmianę komórki:

Warunki ograniczeń

Odwołanie do komórki	Operator	Wartość
<input type="text" value="\$AD\$3"/>	<input "="" type="text" value="<="/>	<input type="text" value="\$AB\$3"/>
<input type="text" value="\$AD\$4"/>	<input "="" type="text" value="="/>	<input type="text" value="\$AB\$4"/>
<input type="text" value="\$AD\$6"/>	<input "="" type="text" value="="/>	<input type="text" value="\$AB\$6"/>
<input type="text" value="\$AD\$7"/>	<input "="" type="text" value="="/>	<input type="text" value="\$AB\$7"/>

Opcje... Pomoc Zamknij **Rozwiąż**

As a result names of selected players were gathered in table with positions they play.

AE	AF	AG	AH
	squad	position	
	Gareth McAuley	DEF	
	Charlie Daniels	DEF	
	Tom Heaton	GK	
	Jermain Defoe	AT	

Table below presents final team members selected by the Solver:

Name	market value[mln]	position	fpl score
Gareth McAuley	1,0	DEF	131
Charlie Daniels	3,0	DEF	134
Tom Heaton	4,0	GK	149
Jermain Defoe	5,0	AT	166
Chris Brunt	4,0	MID	99
Joshua King	8,0	AT	178
Ashley Williams	8,0	DEF	127
Antoine Capoue	9,0	MID	131
James Milner	12,0	MID	139
Gary Cahill	16,0	DEF	178
Georginio Wijnaldum	28,0	MID	149
total	98,0		1581

As we can see the budget is kept and all positions in team are occupied so the Solver did its job very well. By doubling our budget we get 1812 points in total and while multiplying it three times 1940 so the growths aren't that significant. In conclusion maybe it's better to complete team within lower budget and invest saved money in training process to increase players performance as a team.

P.S

I hope that configurations and formulas in cells won't be invalid because of transition between different formats and editors (I used LibreOffice in Polish language version).