FIFA Players Data Analysis Report

09-06-2021

Looking for market opportunities

This report takes data scrapped from www.sofifa.com of players and analyses the data looking for the best opportunities to buy as if you were a football club manager looking for the optimal players, variables considered:

- Age = Players ages.
- Wage = Players wages.
- Best Position = Players best playing position.
- Overall = Players overall skill score.
- Value = Players market values.
- Potential = Players maximum overall potential.

This is an automated and dynamic report, since the data can be scrapped any time from sofifa to update it, and the importance given to each variable can be changed according to the desired of the "manager", for this exercise the importance given for each variable were the following per position:

- Age = 15%
- Wage = 15%
- Overall = 35%
- Value = 30%
- Potential = 5%

With this weight of each variable an optimal player means to have the lowest Age, Wage and Value, with the highest Overall and Potential, giving more importance on Overall and Value.

Index

Data Exploration

General Exploration

Once the dataset was obtained a data wrangling and exploration was made. Too old and too low skill players were filter out, and the numerical data was fixed since it came in the format of $1M or $1.5M or $3K, the $ and letters were removed in order to have int type data to be able to do analysis.

The relationship between players Value and Wage, Overall and Value, were visualized with scatter plots.

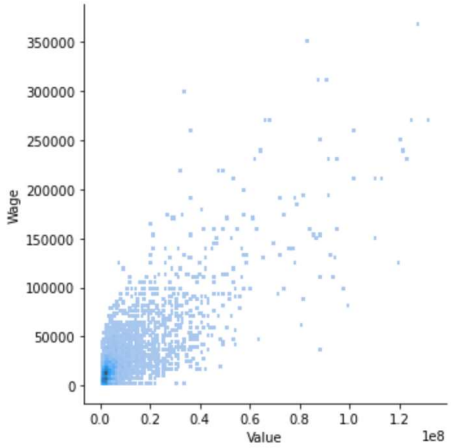

Fig 1. Scatterplot of Value vs Wage

We can see that this looks like a randomized relation, and that **we can get high value players for lower wages than its peers**, those are the type of players we have to look for.
Then we have the visual relationships between Overall and Value.

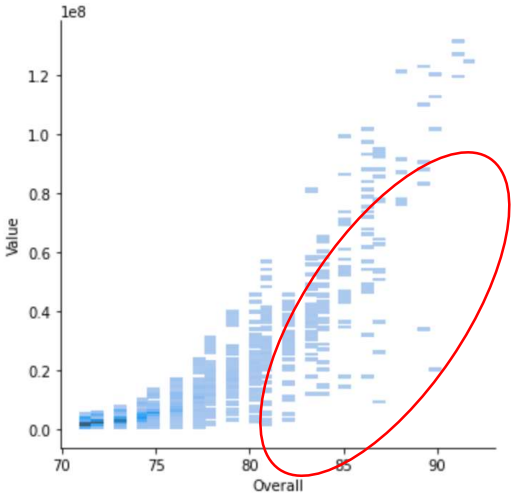

Fig 2. Scatterplots of Value vs Overall. "Sweet spot" marked under red circle.

Here we can see there are player that have lower Value for the same Overall. So, the "sweet spots" of the players we are looking for are marked inside the red circle.

Boxplots

Here we can see a box plot of the Value vs Age of the players, the graph is ordered according to the average value, the highest value average per age is on the left, while the lowest on the right.

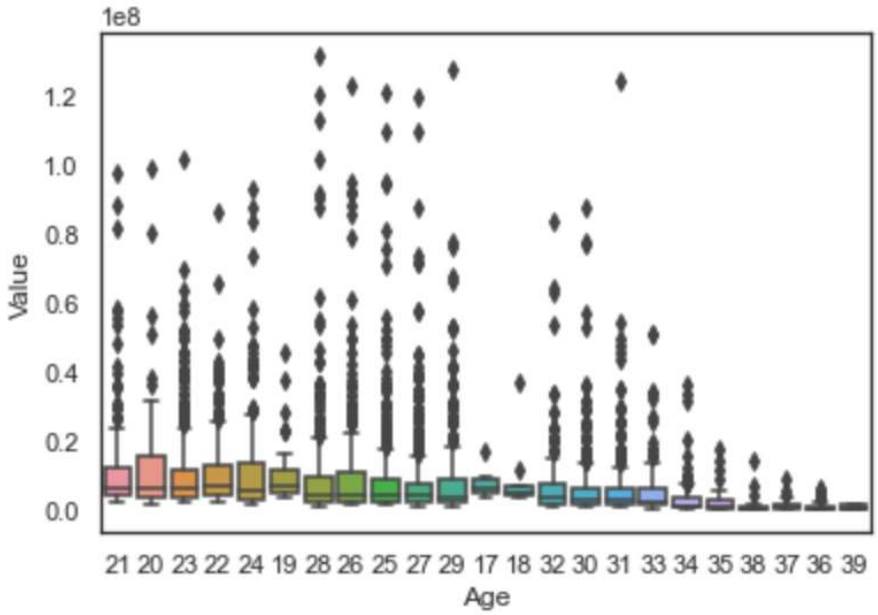

Fig 3. Boxplot Value per Age

Here we can see that the most valued players are 21, 22, 23, 24 years of age, while the lowest valued are over 30 years. We can find a "Sweet spot" for players of 17 and 18 years old whom seems to be low value players and could have good potential, while 21 and 20 years old players could be overvalued, **for best cost-benefit players, 17 and 18 years old seems the best options.**

Now we a box plot of the playing positions of the players vs their value to see how the positions are valued.

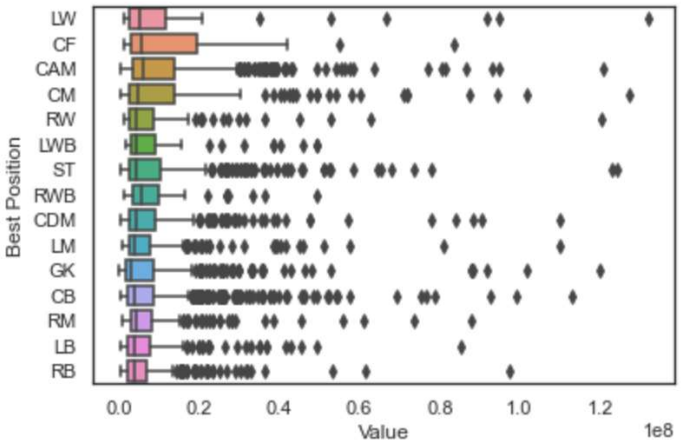

Fig 4. Boxplot Value per Best Position

The positions are ordered form higher average value to lowest from top to bottom. Here we can see that LW, CF, CAM and CM are the most valued positions while GK, CB, RM, LB, RB are the lowest valued

positions. With this we can make better expectations of the difference in values we are going to see per position. **For best cost-benefit outcome we could prefer to buy LMs and RMs over LWs and RWs which are similar positions. Also, RBs and LBs over RWBs and LWBs. As well STs over CFs.**

Then, we see the relation of Best Position and Wage, to see how much players are paid per position, and compare between positions. As in fig. 4 the wages averages are ordered from highest to lowest, top to bottom.

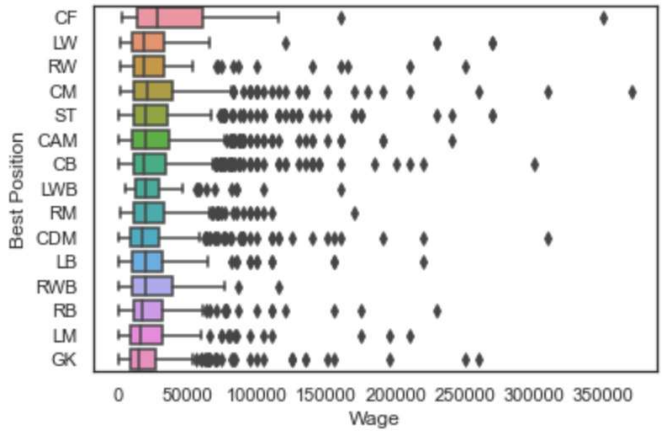

Fig 5. Boxplot Wage per Best Position

Here we can see that CFs are the highest paid position by far, while GK are the lowest. This boxplot shows insides very similar to fig.4 **We should prefer LMs, RMs, STs and LBs over LWs, RWs, CFs and LWBs for lower wages.**

Solution Approach

This was considered as an optimization problem where we have to find the optimal players for each position based on Overall, Wage, Value, Potential and Age.

Different weight was given to each variable, then an optimal index is obtained for each player for each position based on the mean value for that variable, meaning we will find clear outliers for each variable, the players with the lower index are meant to be the optimal players.

More About the index (Technical)

As stated before, the final index is an index obtained using a weighted values technique. The first step was to normalize the data. Then each variable was given a different weight, then all data was put together into the index using the normalize data and the defined weights, with all these variables into account, **the final meaning of the index is that the lower the index the more optimal the player is**.

Normalizing the data

For normalizing the data its distributions were observed, the Age, Overall and Potential data is normal, while the Wage and Value data resembles more of a poisson distribution (since the data is discrete). Here are some of the visualizations.

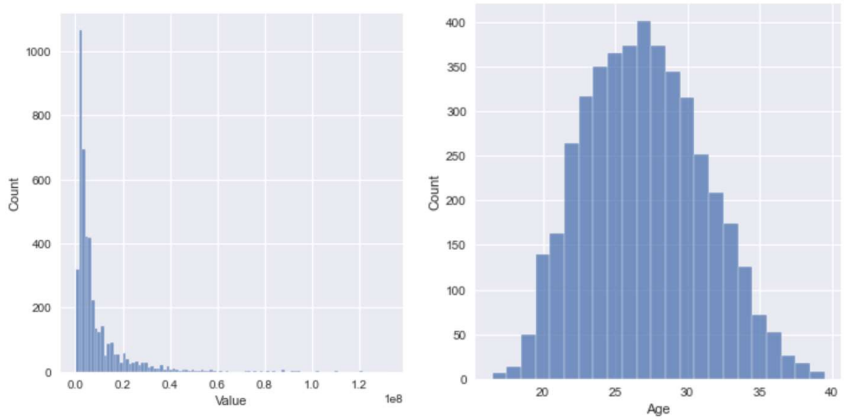

Fig 6. Distributions of Value and Age data

Taking this information into account the data of each variable was normalized using the mean of each variable, the poisson data was adjusted to have similar values to the normal data once everything is normalized.

Variables Weights

Now with data normalized, the variables can be weighted into a single index depending on what importance the user wants to give o each variable. As stated at the beginning the weight given for each variable were the following.

- Age = 15%
- Wage = 15%
- Overall = 35%
- Value = 30%
- Potential = 5%

Results: Optimal Players.

Here are the Top 5 optimal players per position according to the method explained. Pos is the players position, Ove is the Overall skill score, Pot is the potential of the Ove that the player has.

For better data analysis and visualization, the players were stratified with Overall bins, here are the most optimal players for the skill bin of 80 to 100.

Pos	Name	Age	Team	Ove	Pot	Value	Wage	Final Index
ST	A. Isak	20	Real Sociedad	80	85	€ 31,500,000	€ 29,000	2.3648
ST	J. Martinez	27	Atlanta United	80	80	€ 23,000,000	€ 13,000	2.4107
ST	M. Gamez	23	Valencia CF	80	85	€ 32,000,000	€ 40,000	2.4506
ST	S. Haller	26	Ajax	80	83	€ 28,000,000	€ 23,000	2.4513
ST	C. Bakambu	29	Beijing Sinobo FC	80	80	€ 21,000,000	€ 30,000	2.4659

CF	Gabriel Barbosa	22	Inter	81	85	€	23,500,000	€	71,000	2.3850
CF	R. Keane	35	LA Galaxy	80	80	€	5,500,000	€	10,000	2.4191
CF	Joao Pedro	28	Cagliari	80	80	€	21,500,000	€	46,000	2.4617
CF	W. Zaha	27	Crystal Palace	82	82	€	33,000,000	€	82,000	2.5580
CF	H. Mkhitaryan	31	Roma	82	82	€	25,500,000	€	75,000	2.5664
LW	Hulk	33	Shanghai SIPG FC	80	80	€	13,000,000	€	20,000	2.4453
LW	D. Tadic	31	Ajax	84	84	€	35,000,000	€	32,000	2.4474
LW	L. Insigne	29	Napoli	85	85	€	53,000,000	€	120,000	2.5481
LW	R. Sterling	25	Manchester City	87	88	€	95,000,000	€	230,000	2.7196
LW	E. Hazard	29	Real Madrid	86	86	€	66,500,000	€	270,000	2.7266
RW	A. Januzaj	25	Real Sociedad	80	80	€	23,500,000	€	37,000	2.3890
RW	S. Berghuis	28	Feyenoord	81	81	€	26,000,000	€	23,000	2.4344
RW	F. Thauvin	27	Olympique de Marseille	81	81	€	27,500,000	€	54,000	2.4670
RW	C. Vela	31	Los Angeles FC	83	83	€	30,000,000	€	17,000	2.4872
RW	Portu	28	Real Sociedad	83	83	€	36,500,000	€	50,000	2.5162
CAM	Laure Santeiro	20	Fluminense	80	80	€	24,000,000	€	23,000	2.3140
CAM	V. Tsygankov	22	Dynamo Kyiv	80	86	€	34,000,000	€	1,000	2.4069
CAM	David Neres	23	Ajax	80	84	€	30,000,000	€	21,000	2.4347
CAM	Everton	24	SL Benfica	80	83	€	29,000,000	€	16,000	2.4371
CAM	Nuno Padrenda	24	Internacional	81	81	€	29,500,000	€	33,000	2.4535
CM	S. Gerrard	36	LA Galaxy	81	81	€	1,400,000	€	10,000	2.4148
CM	Sergio Oliveira	28	FC Porto	80	80	€	21,000,000	€	19,000	2.4213
CM	T. Savanier	28	Montpellier HSC	80	80	€	21,000,000	€	42,000	2.4572
CM	C. Tolisso	25	FC Bayern Munchen	80	80	€	23,000,000	€	74,000	2.4584
CM	M. Arnold	26	VfL Wolfsburg	81	82	€	27,000,000	€	52,000	2.4623
CDM	Rosberto Dourado	20	Corinthians	81	81	€	25,500,000	€	23,000	2.3068
CDM	Y. Rakitskyi	28	Shakhtar Donetsk	80	80	€	12,000,000	€	2,000	2.3466
CDM	B. Kamara	20	Olympique de Marseille	80	86	€	30,000,000	€	30,000	2.3863
CDM	M. Locatelli	22	Sassuolo	80	86	€	31,000,000	€	32,000	2.4416
CDM	Danilo Pereira	28	Paris Saint-Germain	81	81	€	22,500,000	€	19,000	2.4486
LM	Ricardo Horta	25	SC Braga	80	81	€	25,000,000	€	19,000	2.4337
LM	Otavio	25	FC Porto	81	82	€	31,000,000	€	19,000	2.4631
LM	Cucurella	21	Getafe CF	81	87	€	41,500,000	€	28,000	2.4777
LM	D. Machs	27	Granada CF	80	80	€	22,500,000	€	33,000	2.4863
LM	Taison	32	Shakhtar Donetsk	81	81	€	20,500,000	€	2,000	2.5126
RM	J. Bamba	24	LOSC Lille	80	83	€	29,000,000	€	38,000	2.4778
RM	E. Salvio	29	Boca Juniors	80	80	€	20,500,000	€	24,000	2.4939
RM	Rafa	27	SL Benfica	83	83	€	38,500,000	€	23,000	2.5377
RM	M. Lazzari	26	Lazio	81	81	€	28,000,000	€	70,000	2.5556
RM	Q. Promes	28	Spartak Moscow	80	80	€	21,000,000	€	68,000	2.5600
LWB	Grimaldo	24	SL Benfica	82	85	€	38,500,000	€	16,000	2.5167
LWB	K. Tierney	23	Arsenal	80	86	€	31,000,000	€	69,000	2.5747
LWB	Jonny	26	Wolverhampton	80	81	€	22,500,000	€	81,000	2.5963
LWB	AngeliÃ±o	23	RB Leipzig	83	86	€	46,000,000	€	58,000	2.6155
LWB	B. Chilwell	23	Chelsea	82	86	€	40,500,000	€	85,000	2.6429

LB	P. Max	26	PSV	80	81	€	22,500,000	€	22,000	2.4393
LB	Ismaily	30	Shakhtar Donetsk	81	81	€	22,000,000	€	1,000	2.4658
LB	Nacho Monreal	34	Real Sociedad	80	80	€	7,500,000	€	32,000	2.5070
LB	Renan Lodi	22	AtlÃ©tico Madrid	81	86	€	37,000,000	€	49,000	2.5171
LB	N. Tagliafico	27	Ajax	83	83	€	35,000,000	€	25,000	2.5237
CB	D. Dumfries	24	PSV	80	83	€	26,500,000	€	19,000	2.4299
CB	Yeray	25	Athletic Club de Bilbao	80	82	€	23,500,000	€	26,000	2.4354
CB	Juiano Mestres	24	AtlÃ©tico Mineiro	81	81	€	26,000,000	€	37,000	2.4370
CB	V. Kompany	33	RSC Anderlecht	82	82	€	10,500,000	€	29,000	2.4614
CB	Marcano	33	FC Porto	80	80	€	10,000,000	€	17,000	2.4646
RB	Maikel Catarino	24	Corinthians	80	80	€	22,000,000	€	26,000	2.4122
RB	R. James	20	Chelsea	80	86	€	30,500,000	€	61,000	2.4867
RB	Ruben Perea	28	Villarreal CF	80	80	€	19,000,000	€	37,000	2.4910
RB	S. Lainer	27	Borussia Monchenglad	80	80	€	20,500,000	€	41,000	2.4911
RB	Capa	28	Athletic Club de Bilbao	81	81	€	23,500,000	€	31,000	2.5084
RWB	K. Laimer	23	RB Leipzig	81	85	€	33,500,000	€	49,000	2.5545
RWB	Nelson Semedo	26	Wolverhampton	81	82	€	27,500,000	€	86,000	2.6387
RWB	Mario Fernandes	29	PFC CSKA Moscow	82	82	€	27,000,000	€	60,000	2.6387
RWB	A. Wan-Bissaka	22	Manchester United	83	87	€	49,500,000	€	115,000	2.7686
GK	Heniton Pires	24	Flamengo	80	80	€	18,000,000	€	28,000	2.3927
GK	Adan	33	Sporting CP	81	81	€	8,500,000	€	14,000	2.4296
GK	S. Mandanda	35	Olympique de Marseille	82	82	€	3,400,000	€	26,000	2.4318
GK	Unai Simon	23	Athletic Club de Bilbao	80	85	€	25,500,000	€	18,000	2.4381
GK	L. Fabianski	35	West Ham United	81	81	€	2,900,000	€	26,000	2.4416