

Football Players Value Prediction

Burak Arslan



Introduction

Is it possible to determine a player's value by his statistics ?

Rk	Player	Nation	Pos	Squad	Age	Born	Total					Short			Medium			Long											
							Cmp	Att	Cmp%	TotDist	PrgDist	Cmp	Att	Cmp%	Cmp	Att	Cmp%	Cmp	Att	Cmp%	Ast	xA	A-xA	KP	1/3	PPA	CrsPA	Prog	
1	Patrick van Aanholt		DF	Crystal Palace	28	1990	27.9	1227	1558	78.8	21251	9280	578	647	89.3	521	634	82.2	112	221	50.7	2	3.2	-1.2	32	86	25	5	129
2	Max Aarons		DF	Norwich City	19	2000	36.0	1316	1684	78.1	20293	6834	733	820	89.4	475	596	79.7	78	174	44.8	1	3.4	-2.4	24	43	32	13	77
3	Tammy Abraham		FW	Chelsea	21	1997	24.6	288	412	69.9	4213	811	159	222	71.6	77	101	76.2	26	35	74.3	3	2.5	+0.5	18	16	5	0	31
4	Che Adams		FW	Sheff Wed	23	1999	24.3	222	322	69.3	3001	593	62	103	60.2	22	33	66.3	12	23	52.2	1	2.3	-1.3	11	12	3	0	12
5	Adrián		DF	Real Betis	29	1990	33.0	1227	1558	78.8	21251	9280	578	647	89.3	521	634	82.2	112	221	50.7	2	3.2	-1.2	32	86	25	5	129
6	Sergio Agüero		FW	Manchester City	34	1988	35.0	1227	1558	78.8	21251	9280	578	647	89.3	521	634	82.2	112	221	50.7	2	3.2	-1.2	32	86	25	5	129
7	Albian Ajeti		FW	Sheff Wed	23	1999	24.3	222	322	69.3	3001	593	62	103	60.2	22	33	66.3	12	23	52.2	1	2.3	-1.3	11	12	3	0	12
8	Nathan Aspinall		DF	Crystal Palace	28	1990	27.9	1227	1558	78.8	21251	9280	578	647	89.3	521	634	82.2	112	221	50.7	2	3.2	-1.2	32	86	25	5	129
9	Marc Albright		DF	Crystal Palace	28	1990	27.9	1227	1558	78.8	21251	9280	578	647	89.3	521	634	82.2	112	221	50.7	2	3.2	-1.2	32	86	25	5	129
10	Tobias Aldrich		DF	Crystal Palace	28	1990	27.9	1227	1558	78.8	21251	9280	578	647	89.3	521	634	82.2	112	221	50.7	2	3.2	-1.2	32	86	25	5	129
11	Trent Alexander-Arnold		DF	Liverpool	25	1998	25.0	1227	1558	78.8	21251	9280	578	647	89.3	521	634	82.2	112	221	50.7	2	3.2	-1.2	32	86	25	5	129
12	Alisson		GK	Liverpool	32	1992	31.0	1227	1558	78.8	21251	9280	578	647	89.3	521	634	82.2	112	221	50.7	2	3.2	-1.2	32	86	25	5	129
13	Dele Alli		FW	Sheff Wed	23	1999	24.3	222	322	69.3	3001	593	62	103	60.2	22	33	66.3	12	23	52.2	1	2.3	-1.3	11	12	3	0	12
14	Miquel Almiron		FW	Sheff Wed	23	1999	24.3	222	322	69.3	3001	593	62	103	60.2	22	33	66.3	12	23	52.2	1	2.3	-1.3	11	12	3	0	12
15	Marcos Alonso		FW	Sheff Wed	23	1999	24.3	222	322	69.3	3001	593	62	103	60.2	22	33	66.3	12	23	52.2	1	2.3	-1.3	11	12	3	0	12
16	Steven Alzola		FW	Sheff Wed	23	1999	24.3	222	322	69.3	3001	593	62	103	60.2	22	33	66.3	12	23	52.2	1	2.3	-1.3	11	12	3	0	12
17	Ibrahim Amara		FW	Sheff Wed	23	1999	24.3	222	322	69.3	3001	593	62	103	60.2	22	33	66.3	12	23	52.2	1	2.3	-1.3	11	12	3	0	12
18	Felipe Anderson		FW	Sheff Wed	23	1999	24.3	222	322	69.3	3001	593	62	103	60.2	22	33	66.3	12	23	52.2	1	2.3	-1.3	11	12	3	0	12
19	Florin Andon		FW	Sheff Wed	23	1999	24.3	222	322	69.3	3001	593	62	103	60.2	22	33	66.3	12	23	52.2	1	2.3	-1.3	11	12	3	0	12
20	Angelito		FW	Sheff Wed	23	1999	24.3	222	322	69.3	3001	593	62	103	60.2	22	33	66.3	12	23	52.2	1	2.3	-1.3	11	12	3	0	12
21	Faustino Aspas		FW	Sheff Wed	23	1999	24.3	222	322	69.3	3001	593	62	103	60.2	22	33	66.3	12	23	52.2	1	2.3	-1.3	11	12	3	0	12
22	Michael Awe		FW	Sheff Wed	23	1999	24.3	222	322	69.3	3001	593	62	103	60.2	22	33	66.3	12	23	52.2	1	2.3	-1.3	11	12	3	0	12
23	Stuart Armstrong		FW	Sheff Wed	23	1999	24.3	222	322	69.3	3001	593	62	103	60.2	22	33	66.3	12	23	52.2	1	2.3	-1.3	11	12	3	0	12
24	Kean Aspinall		DF	Crystal Palace	28	1990	27.9	1227	1558	78.8	21251	9280	578	647	89.3	521	634	82.2	112	221	50.7	2	3.2	-1.2	32	86	25	5	129
25	Christian Aspinall		DF	Crystal Palace	28	1990	27.9	1227	1558	78.8	21251	9280	578	647	89.3	521	634	82.2	112	221	50.7	2	3.2	-1.2	32	86	25	5	129
Rk	Player	Nation	Pos	Squad	Comp	Age	Born	MP	Min	MP/MP	Min% 90s	Starts	Mn/Start	Compl	Subs	Mn/Sub	unSub	PPM	onG	onGA	+/-	+/-90	On-Off	onxG	onxGA	xG+/-	xG+/-90	On-Off	
1	Patrick van Aanholt		DF	Crystal Palace		28	1990	29	2,507	86	73.3	27.9	29	86	25	0	0	0	1.24	21	38	-17	-0.61	-0.41	26.6	37.2	-10.6	-0.38	+0.28
2	Max Aarons		DF	Norwich City		19	2000	36	3,240	36.0	0	0	0	0	0	0	0	0	0.00	0.03	0.03	0.00	0.03	0.05	0.3	0.4	0.01	0.09	0.11
3	Yunis Abdelhamid		DF	Reims		31	1987	28	2,520	28.0	0	0	0	0	0	0	0	0	0.11	0.00	0.11	0.11	0.11	1.3	1.3	0.1	0.05	0.00	0.11
4	Suleiman Abdullahi		FW,MF	Union Berlin		22	1996	6	1	157	1.7	1	0	0	0	0	0	0	0.57	0.00	0.57	0.57	0.57	1.3	1.3	0.2	0.72	0.10	0.82
5	Mehdi Abeid		MF	Nantes		26	1992	25	20	1,817	20.2	1	2	0	0	0	0	0	0.05	0.10	0.15	0.05	0.15	1.6	1.6	1.4	0.08	0.07	0.15
6	Charles Abi		FW	Saint-Étienne		19	2000	15	6	623	6.9	0	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.2	0.2	1.1	0.03	0.16	0.19
7	David Abraham		DF	Eint Frankfurt		33	1986	20	19	1,688	18.8	2	0	0	0	0	0	0	0.11	0.00	0.11	0.11	0.11	1.4	1.4	0.1	0.07	0.00	0.08
8	Tammy Abraham		FW	Chelsea		21	1997	34	25	2,215	24.6	15	3	0	0	0	0	0	0.61	0.12	0.73	0.61	0.73	14.1	14.1	2.5	0.57	0.10	0.67
9	Amir Abrashi		MF	Freiburg		29	1990	13	8	663	7.4	0	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.4	0.4	0.1	0.05	0.02	0.07
10	Francesco Acerbi		DF	Lazio		31	1988	36	36	3,162	35.1	2	2	0	0	0	0	0	0.06	0.06	0.11	0.06	0.11	2.0	2.0	1.6	0.06	0.05	0.10
11	Che Adams		FW	Southampton		23	1996	30	12	1,111	12.3	4	2	0	0	0	0	0	0.32	0.16	0.49	0.32	0.49	4.0	4.0	1.2	0.32	0.10	0.42
12	Tyler Adams		DF,MF	RB Leipzig		20	1999	14	10	953	10.6	0	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.4	0.4	0.5	0.04	0.05	0.09
13	Sargis Adamyan		FW,MF	Hoffenheim		26	1993	15	6	746	8.3	5	1	0	0	0	0	0	0.60	0.12	0.72	0.60	0.72	4.1	4.1	0.9	0.49	0.11	0.60
14	Antonio Adán		GK	Atlético Madrid		32	1987	1	0	25	0.3	0	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.0	0.00	0.00	0.00
15	Bobby Adegbenye		DF	Lazio		20	1999	11	0	188	2.1	1	0	0	0	0	0	0	0.48	0.00	0.48	0.48	0.48	1.7	1.7	0.1	0.82	0.06	0.88
16	Claud Adjaopong		FW	Hellas Verona		21	1998	5	2	231</																			

Premier League FM 2021 Best Players Review, Profiles

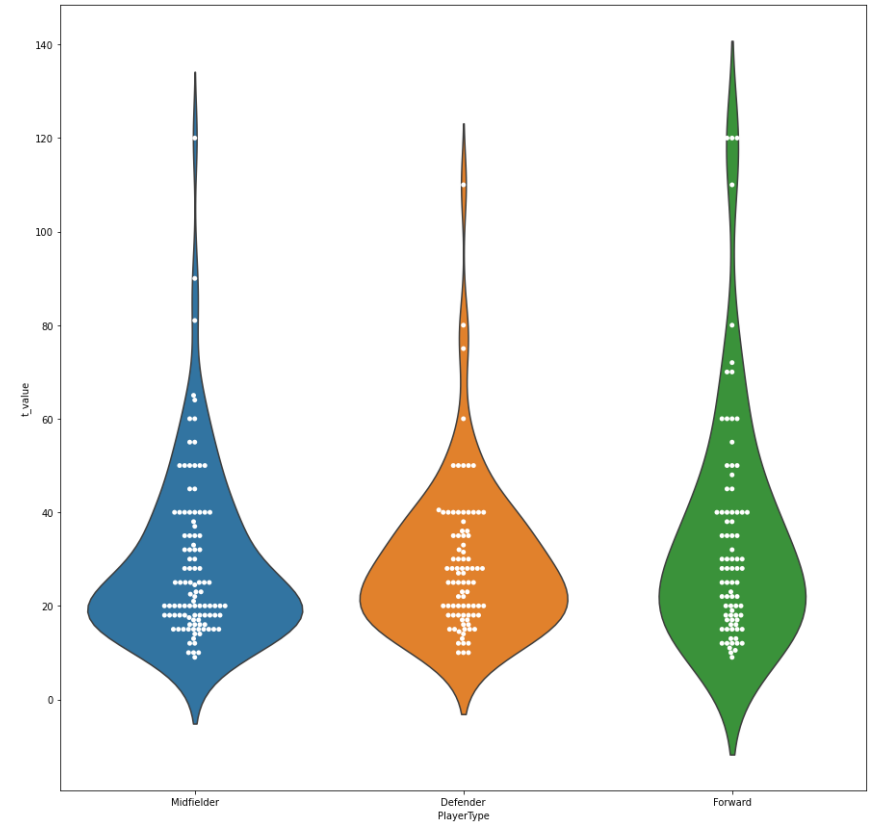
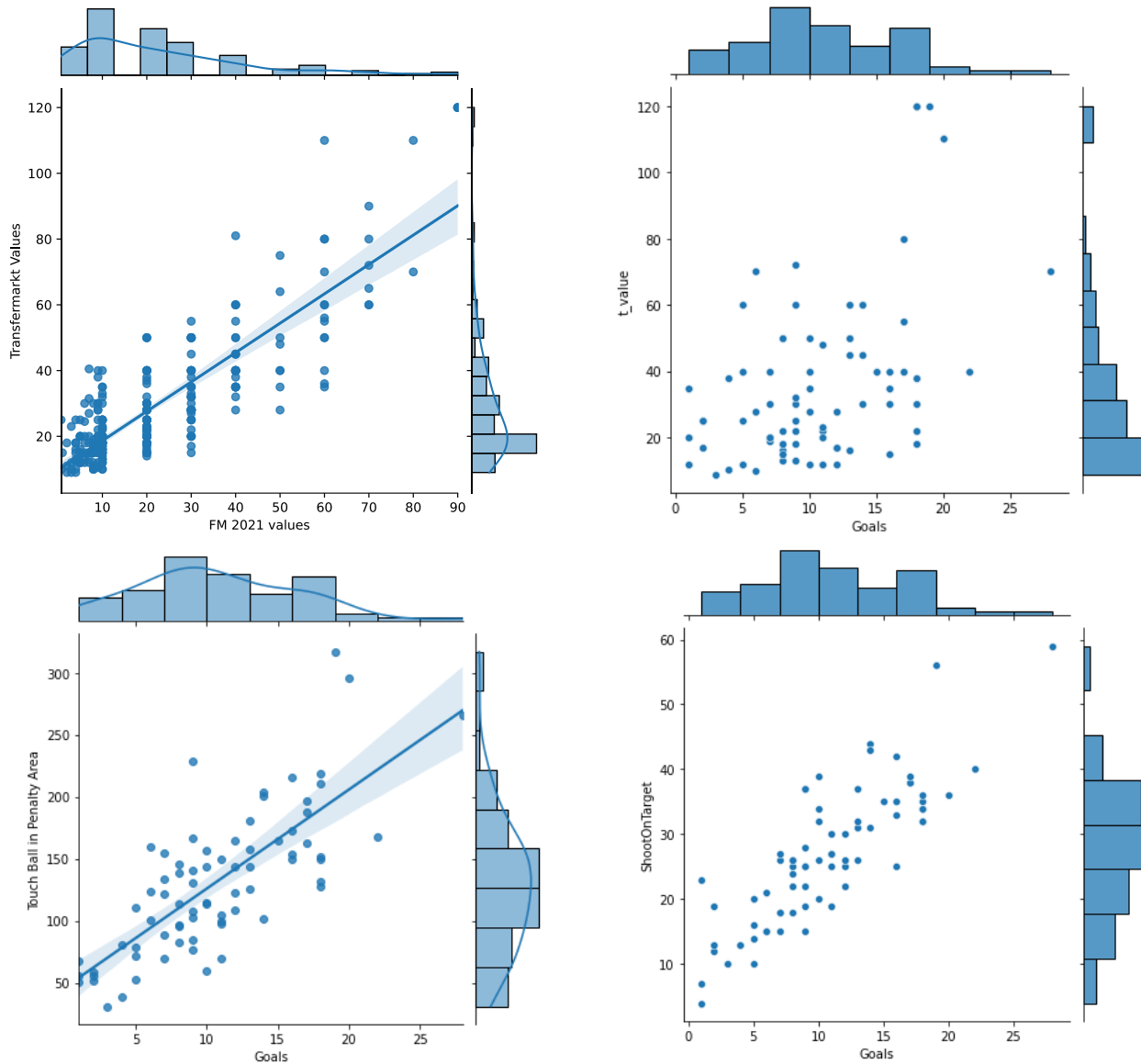
||
||
||

Methodology

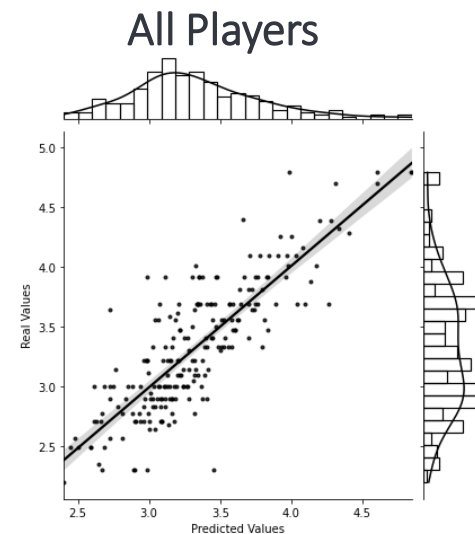
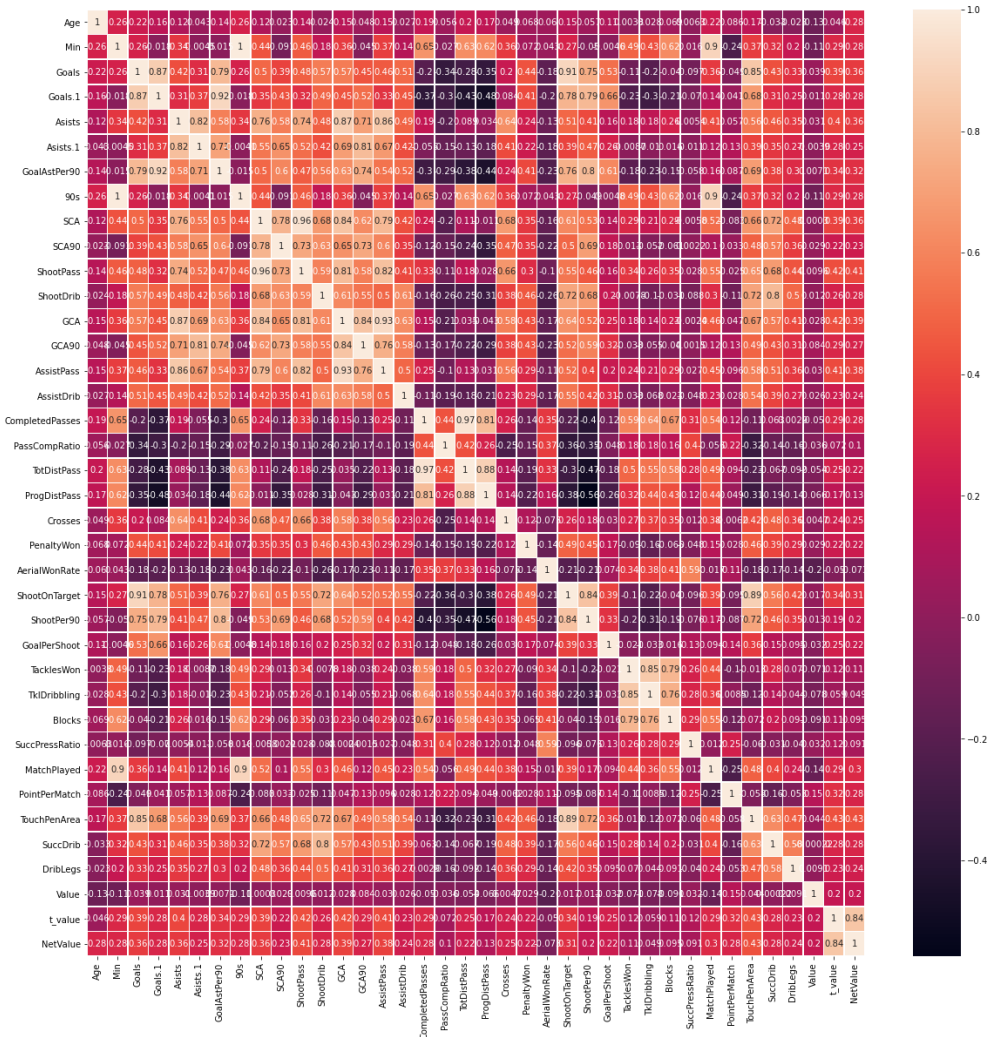
1. Data Fetch & Data process
 - a) Web Scraping
 - I. 2019/2020 Season Statistics
 - II. Football Manager 2021 Game Player Values
 - III. Transfermarkt.de Player Values
2. Exploratory Data Analysis
 - a) Feature Tuning
3. OLS & Regression Models
4. Prediction



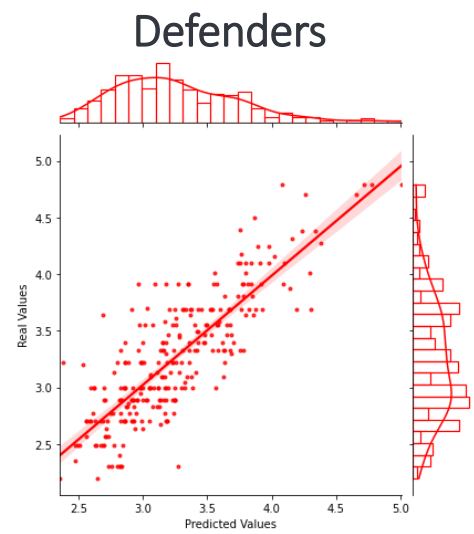
Exploratory Data Analysis



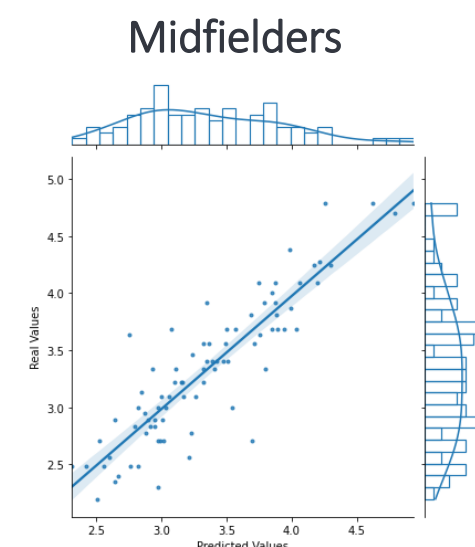
Results



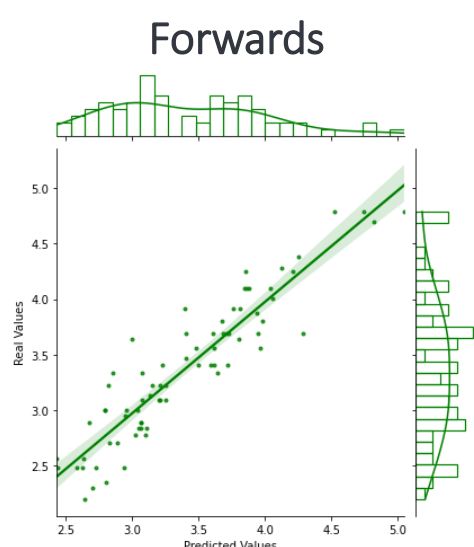
R² score : 0.69



R² score : 0.68



R² score : 0.79



R² score : 0.86

Results

New Features for Better Fitting

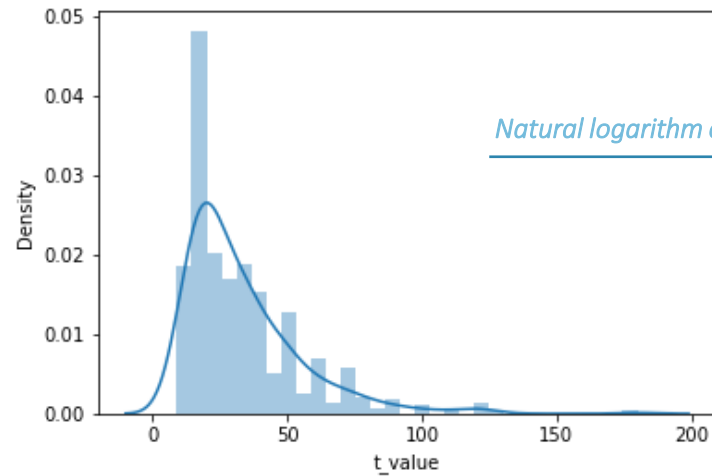
- 1- Agegroup instead of Age
- 2- Winner : (Point per Match Played \wedge 2 * Played Match count)
- 3- Goal Change Creator instead of Assist, Shoot on Target, Dribbling lead to assist, pass lead to assist, Penalty won, Successful pressing at 3rd Area
- 4- Features that has high correlation between each other had been removed.

Features that were converted to categorical variable into dummy/indicator variables :

- 1- League
- 2- Player Position

And many features had been removed to overcome overfitting !

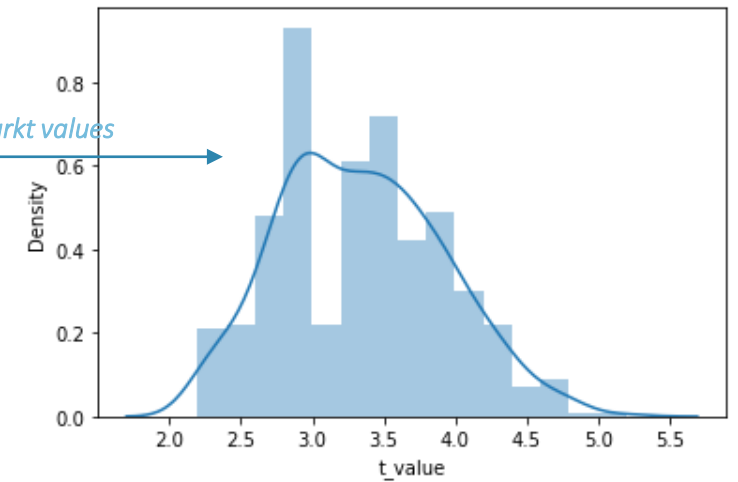
Players that were substitute mostly, also had been removed.



OLS Regression Results

Dep. Variable:	v_value	R-squared (uncentered):	0.865
Model:	OLS	Adj. R-squared (uncentered):	0.847
Method:	Least Squares	F-statistic:	46.68
Date:	Sun, 10 Jan 2021	Prob (F-statistic):	3.95e-91
Time:	13:14:21	Log-Likelihood:	-1157.9
No. Observations:	289	AIC:	2386.
Df Residuals:	254	BIC:	2514.
Df Model:	35		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
Age	-1.1072	0.291	-3.801	0.000	-1.681	-0.534
Min	-0.6481	0.346	-1.874	0.062	-1.329	0.033
Goals	1.6613	0.714	2.326	0.021	0.255	3.068
Goals.1	-141.0973	218.693	-0.645	0.519	-571.780	289.585
Asists	1.5685	1.160	1.352	0.178	-0.716	3.853
Asists.1	-137.4894	218.400	-0.630	0.530	-567.596	292.617
GoalAstPer90	131.2717	218.570	0.601	0.549	-299.168	561.711
90s	58.1142	31.115	1.868	0.063	-3.161	119.390
SCA	-0.1565	0.161	-0.974	0.331	-0.473	0.160
SCA90	0.3174	2.215	0.143	0.886	-4.045	4.680
ShootPass	0.2867	0.196	1.460	0.145	-0.100	0.673



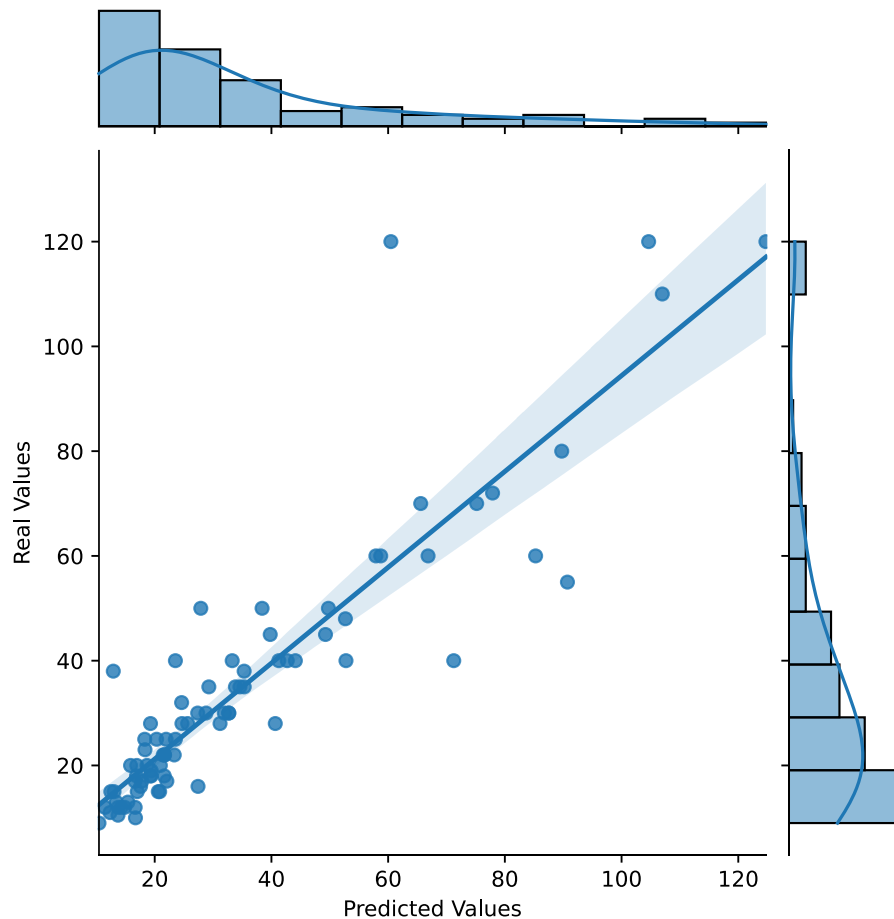
OLS Regression Results

Dep. Variable:	v_value	R-squared (uncentered):	0.984
Model:	OLS	Adj. R-squared (uncentered):	0.982
Method:	Least Squares	F-statistic:	442.1
Date:	Sun, 10 Jan 2021	Prob (F-statistic):	3.79e-207
Time:	13:19:11	Log-Likelihood:	-158.64
No. Observations:	289	AIC:	387.3
Df Residuals:	254	BIC:	515.6
Df Model:	35		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
Age	-0.0106	0.009	-1.158	0.248	-0.029	0.007
Min	-0.0217	0.011	-1.996	0.047	-0.043	-0.000
Goals	0.0399	0.022	1.773	0.077	-0.004	0.084
Goals.1	-4.9898	6.890	-0.724	0.470	-18.558	8.579
Asists	0.0098	0.037	0.269	0.788	-0.062	0.082
Asists.1	-4.4356	6.881	-0.645	0.520	-17.986	9.115
GoalAstPer90	4.5283	6.886	0.658	0.511	-9.032	18.089
90s	1.9621	0.980	2.002	0.046	0.032	3.893
SCA	0.0010	0.005	0.206	0.837	-0.009	0.011

Conclusions

- For more accuracy, models must build on each player type.
- For more good results, we need more data, not more features.
- The statistics is not a fair enough to determine players value.



PlayerName	Transfermarkt Value	Prediction	Difference
Mohamed Salah	120.0	155.318826	35.318826
Harry Kane	120.0	92.070160	-27.929840
Sadio Mané	120.0	114.180262	-5.819738
Raheem Sterling	110.0	123.729914	13.729914
Marcus Rashford	80.0	70.329963	-9.670037
Roberto Firmino	72.0	61.710520	-10.289480
Bernardo Silva	70.0	46.966574	-23.033426
Richarlison	60.0	46.549537	-13.450463
Christian Pulisic	60.0	48.336902	-11.663098
Phil Foden	60.0	56.552633	-3.447367
Gabriel Jesus	60.0	47.389728	-12.610272
Anthony Martial	55.0	57.795870	2.795870
Mason Greenwood	50.0	45.297065	-4.702935
Jack Grealish	50.0	43.081864	-6.918136
Riyad Mahrez	48.0	51.430697	3.430697
Dominic Calvert-Lewin	45.0	53.338864	8.338864
Nicolas Pépé	40.0	36.905394	-3.094606
Pierre-Emerick Aubameyang	40.0	39.929488	-0.070512
Diogo Jota	40.0	41.584178	1.584178
Raúl Jiménez	40.0	72.393873	32.393873
Tammy Abraham	40.0	51.913139	11.913139
Sergio Agüero	35.0	37.190653	2.190653
Alexandre Lacazette	35.0	32.565067	-2.434933
Callum Hudson-Odoi	35.0	52.801773	17.801773
Sébastien Haller	30.0	37.046302	7.046302
Moise Kean	25.0	25.875327	0.875327

Future Work

Data Enrichment Options & Challenges :

- Substitute players : Missing Stats.
- Team members may affect each other's statistics.
- Prone to injury.
- Last year on contract.
- Player home country popularity.
- Can play big matches.
- Player leadership.
- Work rate & ethics.
- Web search popularity.
- The economic strength and prestige of the club.

