Should you buy expensive football players?

Team Nothing But Football
BY
Thanesphol Leerungruang (Jack) 6380628
Natanan Nithiittikrai(Mickey) 6380852

2022

Problem Statement 

We see many stars player in football world being paid heavily to entertain the audiences. Many of these players also have ridiculous inflated transfer fees. For example, Neymar Junior the brazilian football player has been transfer from FC Barcelona to Paris Saint-Germain(PSG) for €222 million Euro which was extremely high fee. We want to see if these values are actually related to player's performance. So this presentation aims to explore that does the player with higher transfer fees do actually perform better with respect to the transfer fee.

Data Source

Mainly we will be using 2 dataset to find out the correlation between Player's Transfer fee and performance.

1. Player's Transfer fee dataset
   https://www.kaggle.com/code/stmy456/football-transfer-analysis/data?select=top250-00-19.csv
2. Player's Overall Rating dataset
   https://www.kaggle.com/code/oguzkhan/fifa-players-ratings-with-plotly/data?select=fifa_cleaned.csv

01 BIG Cleaning

	Name	Position	Transfer_fee
4211	Neymar	Left Winger	222000000
4455	Kylian Mbappé	Right Winger	135000000
4212	Philippe Coutinho	Attacking Midfield	125000000
4456	Cristiano Ronaldo	Centre-Forward	117000000
4213	Ousmane Dembélé	Right Winger	115000000
...	...	...	...
736	Apostolos Liolidis	Right Winger	900000
735	Sebastián Cejas	Goalkeeper	900000
734	Kléber Pereira	Centre-Forward	900000
740	Cata Díaz	Centre-Back	850000
741	Rémo Meyer	Centre-Back	825000

	name	full_name	positions	overall_rating
0	L. Messi	Lionel Andrés Messi Cuccittini	CF,RW,ST	94
1	C. Eriksen	Christian Dannemann Eriksen	CAM,RM,CM	88
2	P. Pogba	Paul Pogba	CM,CAM	88
3	L. Insigne	Lorenzo Insigne	LW,ST	88
4	K. Koulibaly	Kalidou Koulibaly	CB	88
...	...	...	...	...
17949	R. McKenzie	Rory McKenzie	RM,CAM,CM	67
17950	M. Sipľak	Michal Sipľak	LB	59
17951	J. Bekkema	Jan Bekkema	GK	59
17952	A. Al Yami	Abdulrahman Al Yami	ST,LM	59
17953	Júnior Brumado	José Francisco dos Santos Júnior	ST	59

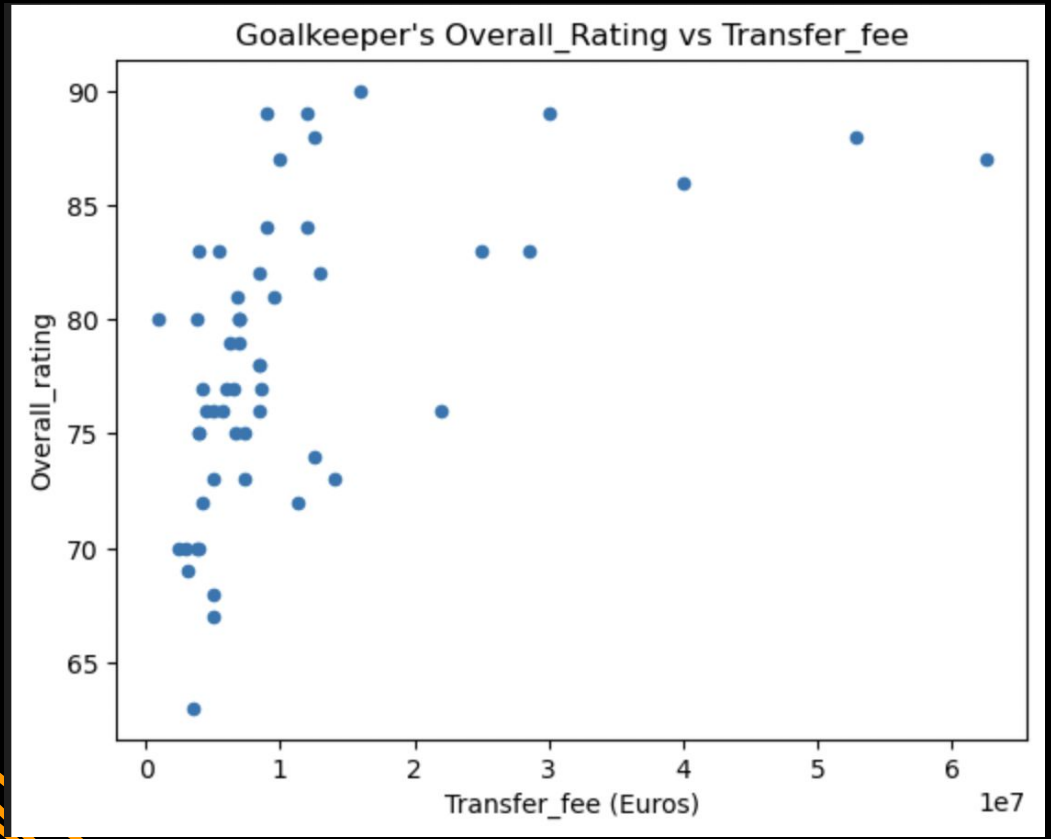
METHODOLOGY

02 Classification

	Name	name	full_name	positions	overall_rating	Transfer_fee
288120	Alisson	Alisson	Alisson Ramses Becker	GK	87	62500000
288261	Gianluigi Buffon	G. Buffon	Gianluigi Buffon	GK	88	52880000
287982	Ederson	Ederson	Ederson Santana de Moraes	GK	86	40000000
3	Manuel Neuer	M. Neuer	Manuel Neuer	GK	89	30000000
286724	Jordan Pickford	J. Pickford	Jordan Pickford	GK	83	28500000
286026	Bernd Leno	B. Leno	Bernd Leno	GK	83	25000000
276508	Alex Meret	A. Meret	Alex Meret	GK	76	22000000
288692	Jan Oblak	J. Oblak	Jan Oblak	GK	90	16000000
259855	Danny Ward	D. Ward	Danny Ward	GK	73	14000000
284637	Jasper Cillessen	J. Cillessen	Jasper Cillessen	GK	82	13000000
288558	Hugo Lloris	H. Lloris	Hugo Lloris	GK	88	12600000
264899	Fraser Forster	F. Forster	Fraser Forster	GK	74	12500000
302	Marc-André ter Stegen	M. ter Stegen	Marc-André ter Stegen	GK	89	12000000
287581	Mattia Perin	M. Perin	Mattia Perin	GK	84	12000000
255945	Angus Gunn	A. Gunn	Angus Gunn	GK	72	11300000
450	Keylor Navas	K. Navas	Keylor Navas	GK	87	10000000
283253	Kevin Trapp	K. Trapp	Kevin Trapp	GK	81	9500000
287175	Yann Sommer	Y. Sommer	Yann Sommer	GK	84	9000000
179	Thibaut Courtois	T. Courtois	Thibaut Courtois	GK	89	8950000
4660	Roberto	Roberto	Roberto Jiménez Gago	GK	77	8600000
274164	Andriy Lunin	A. Lunin	Andriy Lunin	GK	76	8500000
284103	Emiliano Viviano	E. Viviano	Emiliano Viviano	GK	82	8500000
280045	Alban Lafont	A. Lafont	Alban Lafont	GK	78	8500000
279481	Robin Olsen	R. Olsen	Robin Olsen	GK	78	8500000
269272	Sam Johnstone	S. Johnstone	Sam Johnstone	GK	75	7350000
260873	Maarten Stekelenburg	M. Stekelenburg	Maarten Stekelenburg	GK	73	7330000
2018	Diego López	Diego López	Diego López Rodríguez	GK	80	7000000

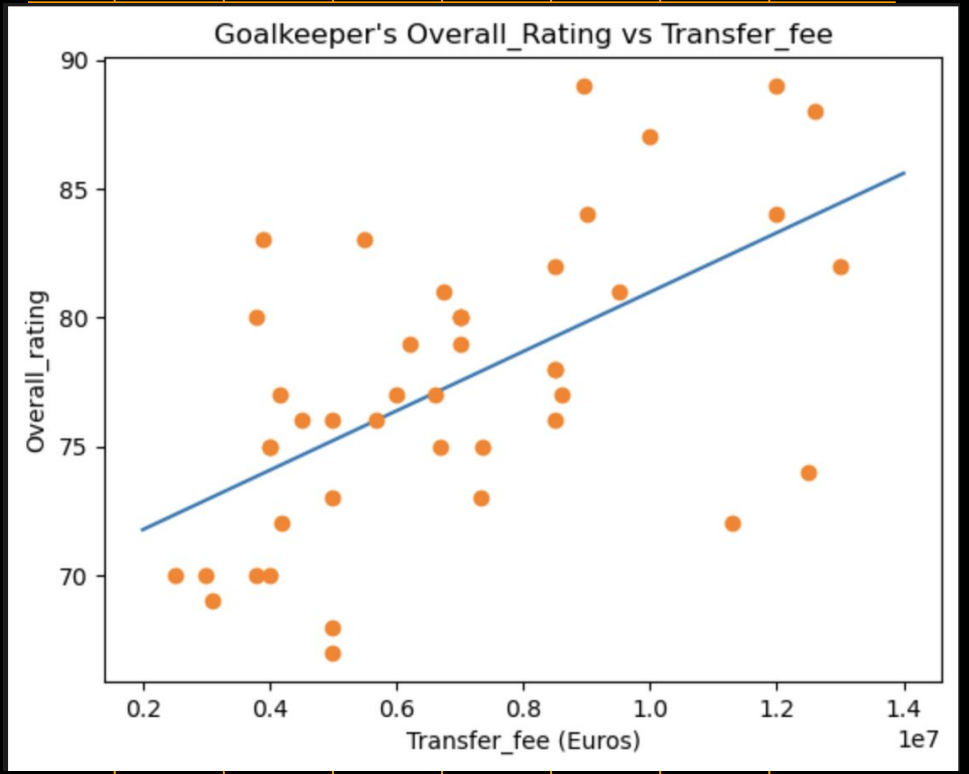
METHODOLOGY

03 Scatter Plot



Plot the graph between Player's overall rating and transfer fee.

METHODOLOGY



Our Methodology

04 Remove Outliers

Remove problematic data for easier model visualization

05 Linear Regression

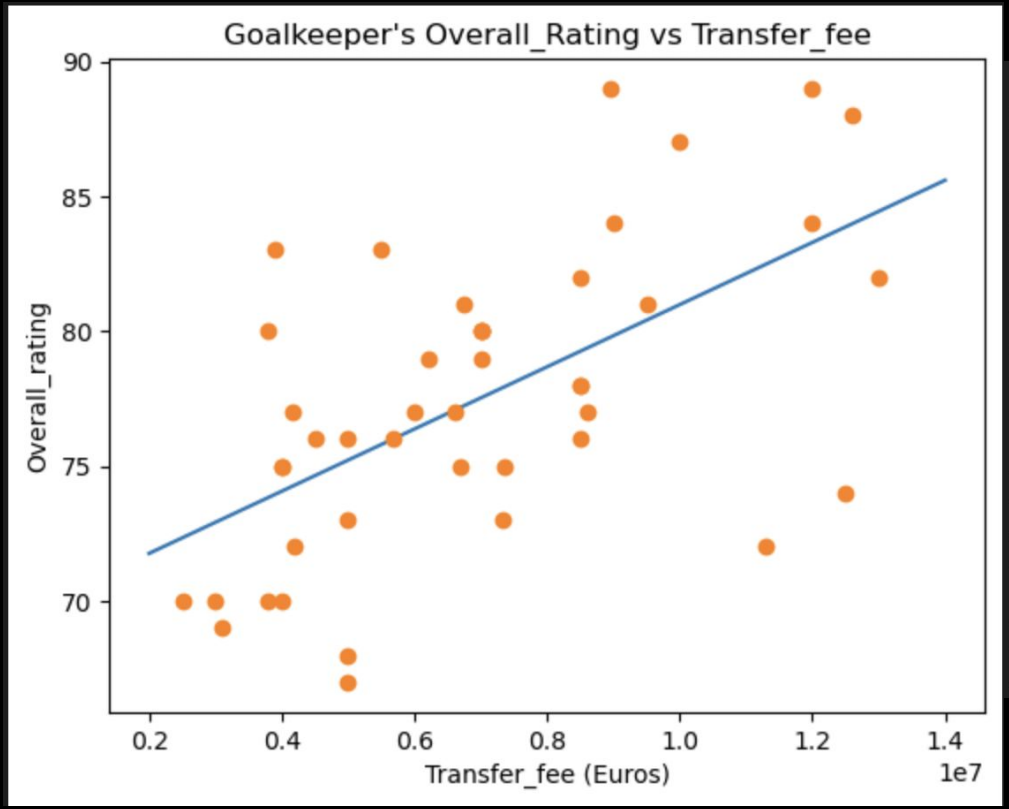
Using what we learnt to find the best fitted line to see the trend of correlation.

06 StatsModel

Using Ordinary Least Square (OLS) to evaluate the uncertainty of the best-fit line, p-value and r-squared.

Visualization

Goalkeeper



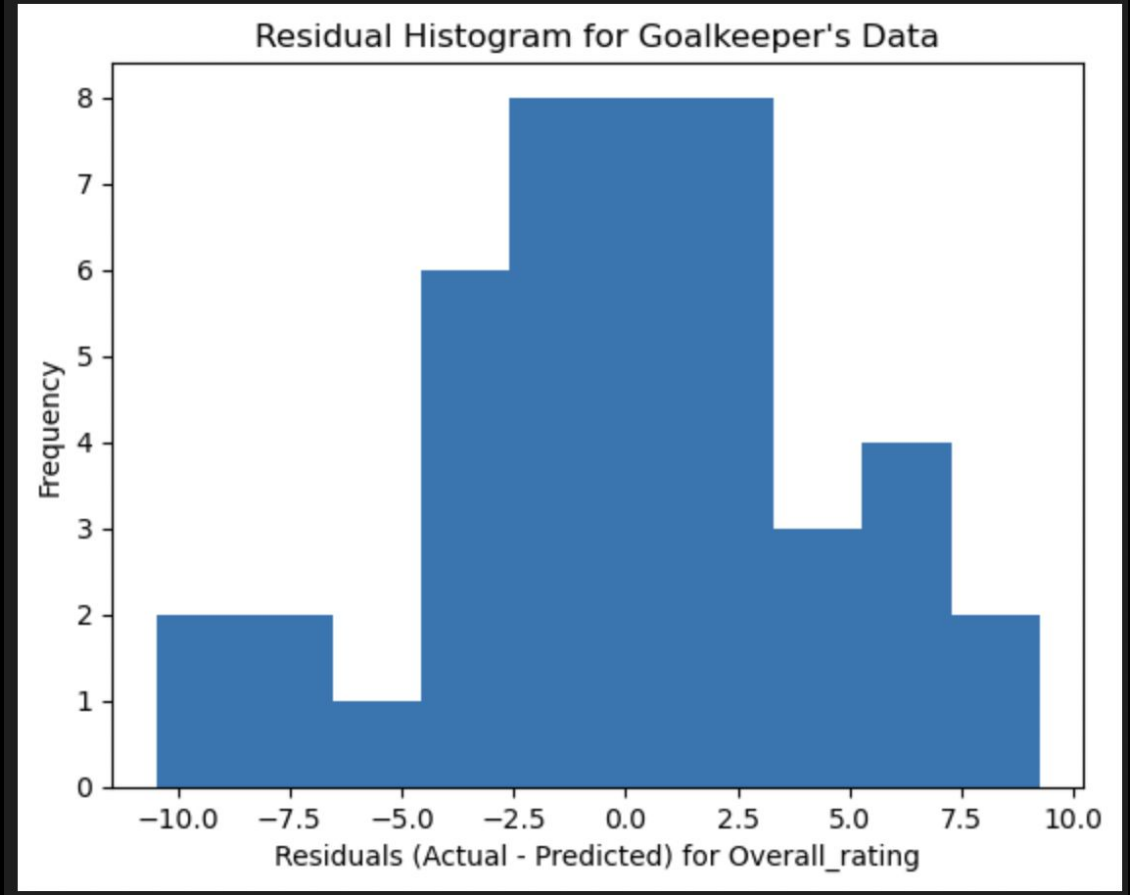
OLS Regression Results

Dep. Variable:	y	R-squared (uncentered):	0.882
Model:	OLS	Adj. R-squared (uncentered):	0.879
Method:	Least Squares	F-statistic:	320.6
Date:	Wed, 30 Nov 2022	Prob (F-statistic):	1.50e-21
Time:	15:45:41	Log-Likelihood:	-206.95
No. Observations:	44	AIC:	415.9
Df Residuals:	43	BIC:	417.7
Df Model:	1		
Covariance Type:	nonrobust		

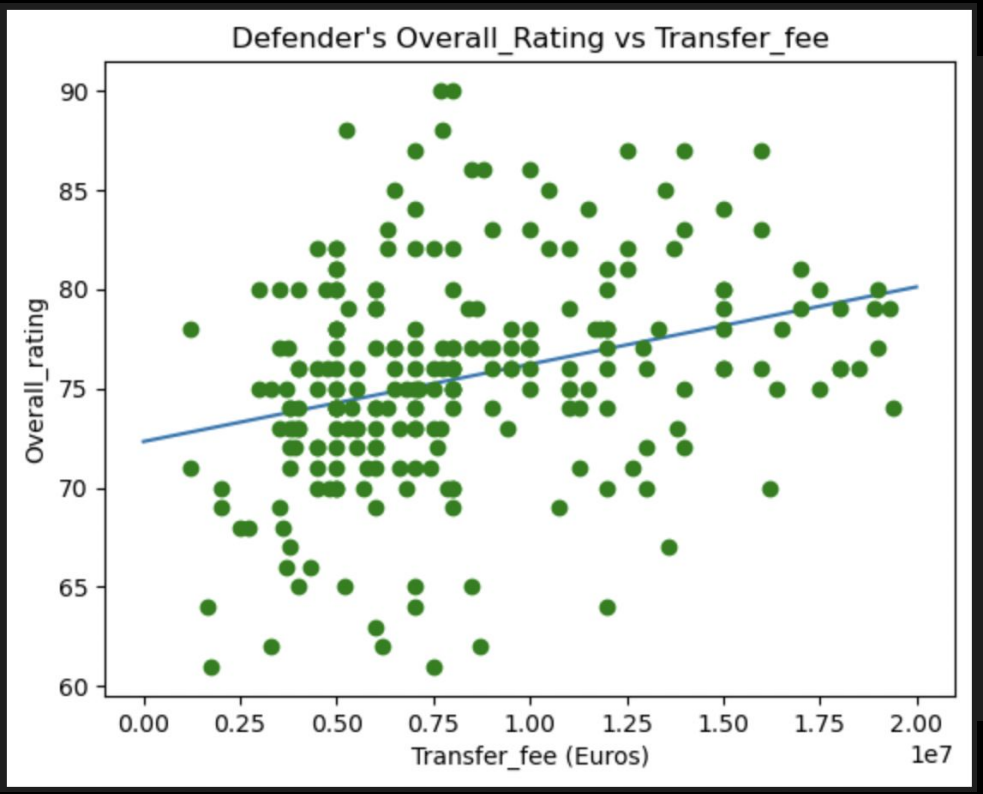
	coef	std err	t	P>|t|	[0.025	0.975]
x1	9.747e-06	5.44e-07	17.905	0.000	8.65e-06	1.08e-05

Omnibus:	3.587	Durbin-Watson:	0.083
Prob(Omnibus):	0.166	Jarque-Bera (JB):	3.365
Skew:	-0.657	Prob(JB):	0.186
Kurtosis:	2.672	Cond. No.	1.00

Goalkeeper (bins n = 10)



Defender



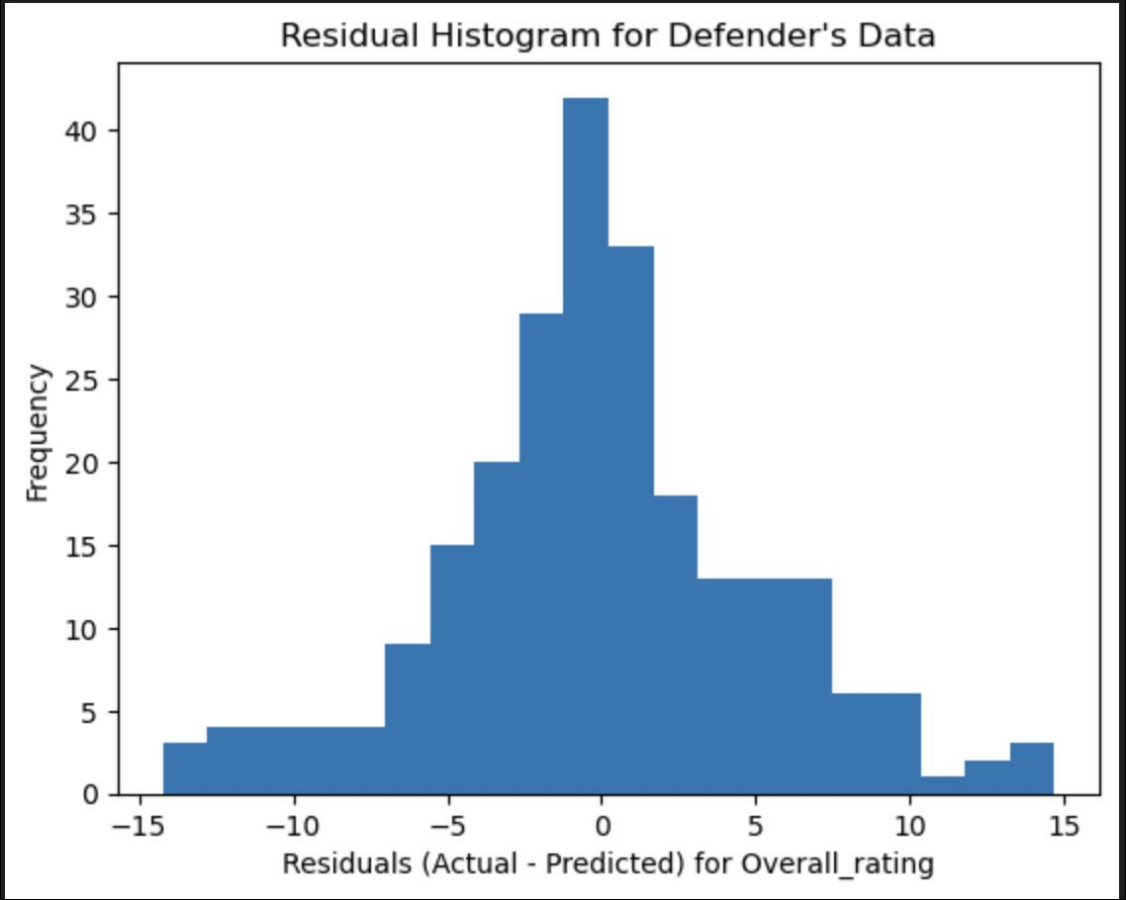
OLS Regression Results

Dep. Variable:	y	R-squared (uncentered):	0.810
Model:	OLS	Adj. R-squared (uncentered):	0.810
Method:	Least Squares	F-statistic:	1030.
Date:	Wed, 30 Nov 2022	Prob (F-statistic):	5.24e-89
Time:	15:45:24	Log-Likelihood:	-1189.3
No. Observations:	242	AIC:	2381.
Df Residuals:	241	BIC:	2384.
Df Model:	1		
Covariance Type:	nonrobust		

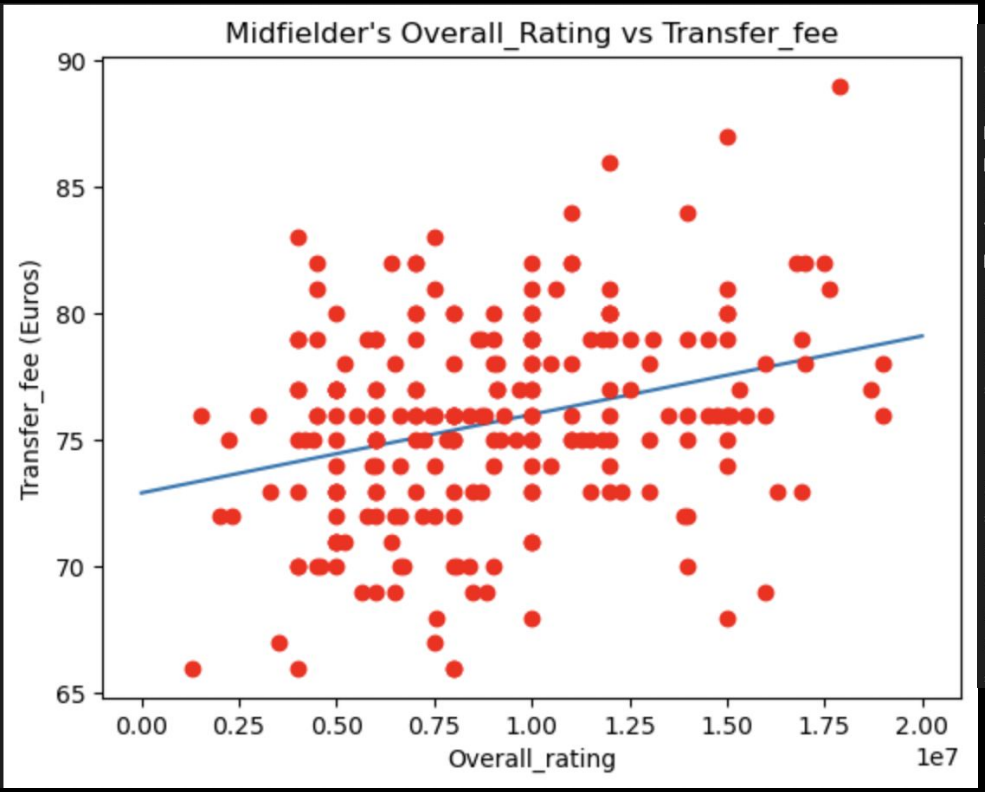
	coef	std err	t	P>|t|	[0.025	0.975]
x1	7.343e-06	2.29e-07	32.101	0.000	6.89e-06	7.79e-06

Omnibus:	24.639	Durbin-Watson:	0.056
Prob(Omnibus):	0.000	Jarque-Bera (JB):	30.007
Skew:	-0.863	Prob(JB):	3.05e-07
Kurtosis:	3.006	Cond. No.	1.00

Defender (bins n = 20)



Midfielder



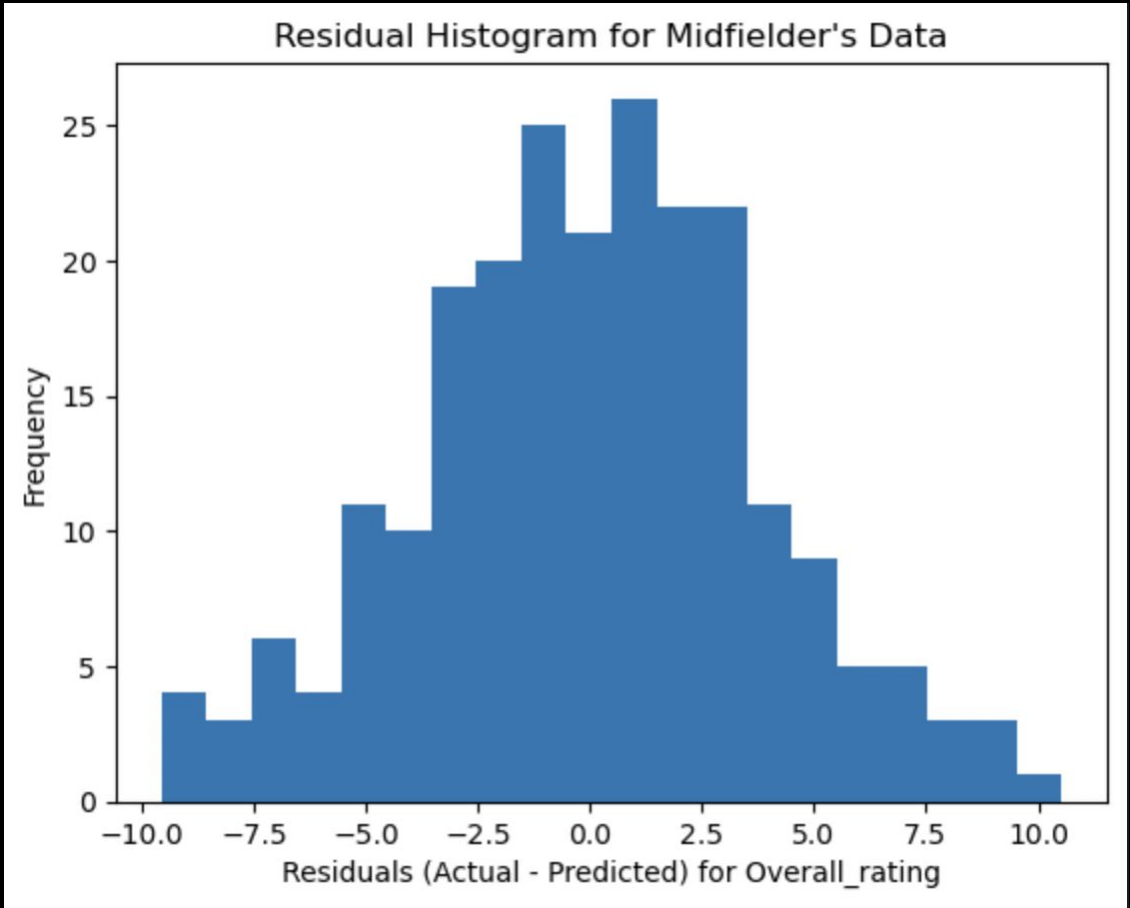
OLS Regression Results

Dep. Variable:	y	R-squared (uncentered):	0.853
Model:	OLS	Adj. R-squared (uncentered):	0.853
Method:	Least Squares	F-statistic:	1332.
Date:	Wed, 30 Nov 2022	Prob (F-statistic):	2.07e-97
Time:	15:44:58	Log-Likelihood:	-1101.2
No. Observations:	230	AIC:	2204.
Df Residuals:	229	BIC:	2208.
Df Model:	1		
Covariance Type:	nonrobust		

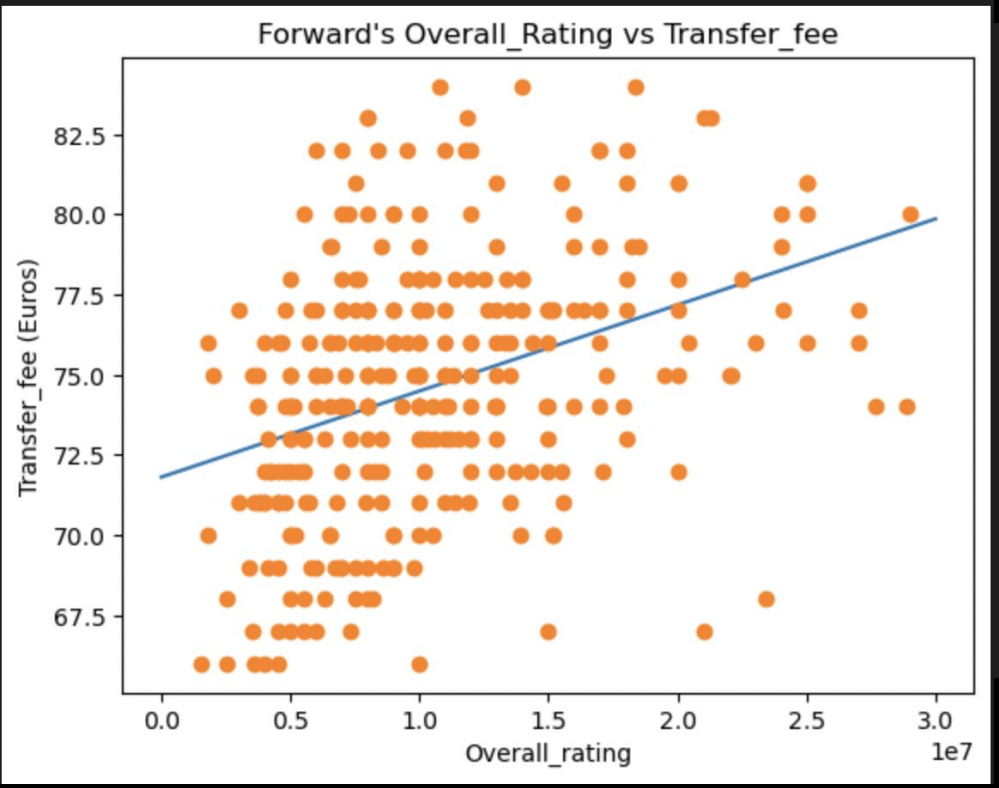
	coef	std err	t	P>|t|	[0.025	0.975]
x1	7.093e-06	1.94e-07	36.495	0.000	6.71e-06	7.48e-06

Omnibus:	11.555	Durbin-Watson:	0.032
Prob(Omnibus):	0.003	Jarque-Bera (JB):	11.277
Skew:	-0.494	Prob(JB):	0.00356
Kurtosis:	2.553	Cond. No.	1.00

Midfielder (bins n = 20)



Forward



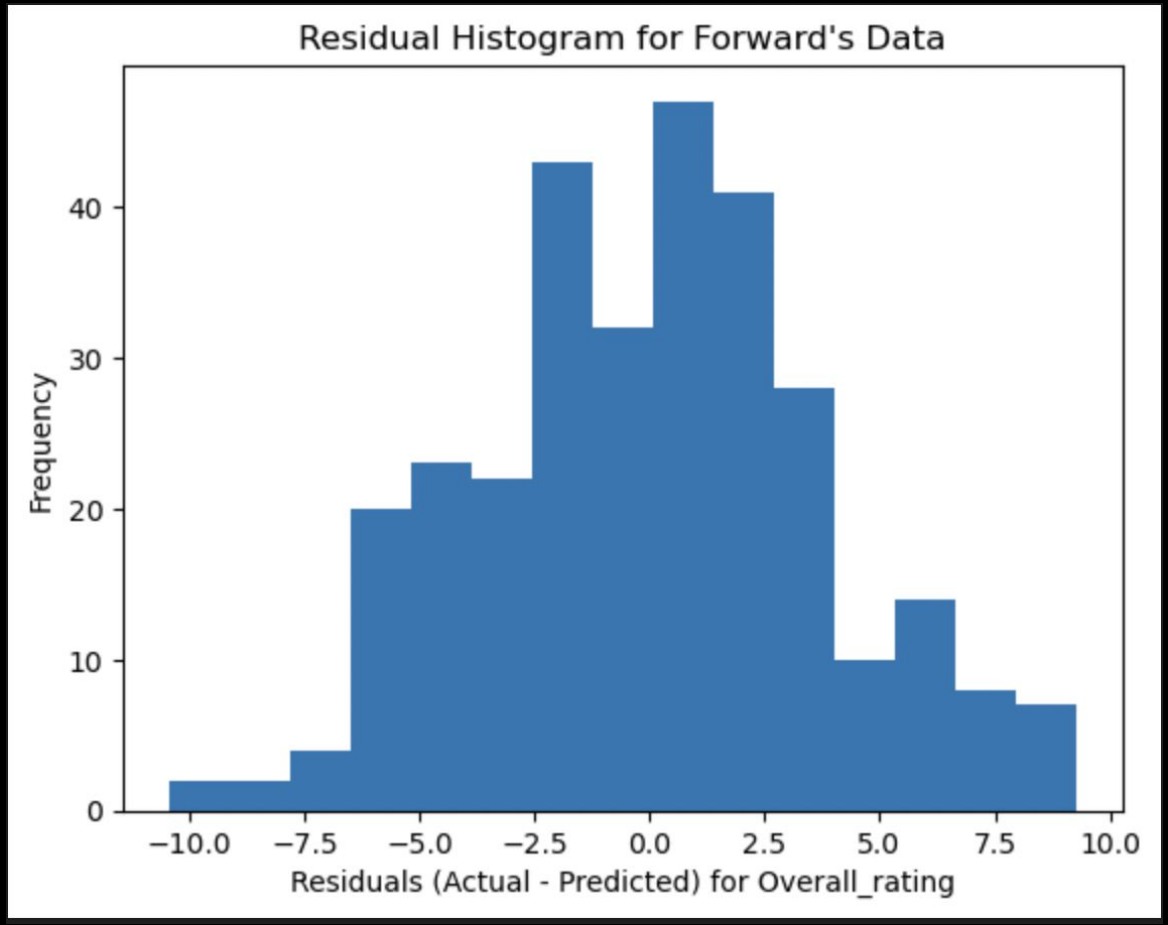
OLS Regression Results

Dep. Variable:	y	R-squared (uncentered):	0.793
Model:	OLS	Adj. R-squared (uncentered):	0.792
Method:	Least Squares	F-statistic:	1155.
Date:	Wed, 30 Nov 2022	Prob (F-statistic):	3.30e-105
Time:	15:37:06	Log-Likelihood:	-1498.8
No. Observations:	303	AIC:	3000.
Df Residuals:	302	BIC:	3003.
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
x1	5.554e-06	1.63e-07	33.983	0.000	5.23e-06	5.88e-06

Omnibus:	46.810	Durbin-Watson:	0.026
Prob(Omnibus):	0.000	Jarque-Bera (JB):	63.916
Skew:	-1.052	Prob(JB):	1.32e-14
Kurtosis:	3.797	Cond. No.	1.00

Forward (bins n = 15)



The Result

As you see from the visualization we can see a slightly positive correlation between the player with a high transfer fee also have high overall rating as well.
In conclusion, it is definitely worth it for you to pay more to get a better football player in FIFA or if you ever become an owner of a football team in the league.


FIFA 22
HOW TO
BUY
PLAYERS
FOR
BEST PRICE
PLAYER NAME
Anthony Martial
PRICING
Price
Any
Max. Price
Any
Min. Buy Now
Any
Max. Buy Now

THANK YOU