



Who Will be the Next Football Superstar?

– Using FIFA game’s data to predict player’s potential

Siyan Peng

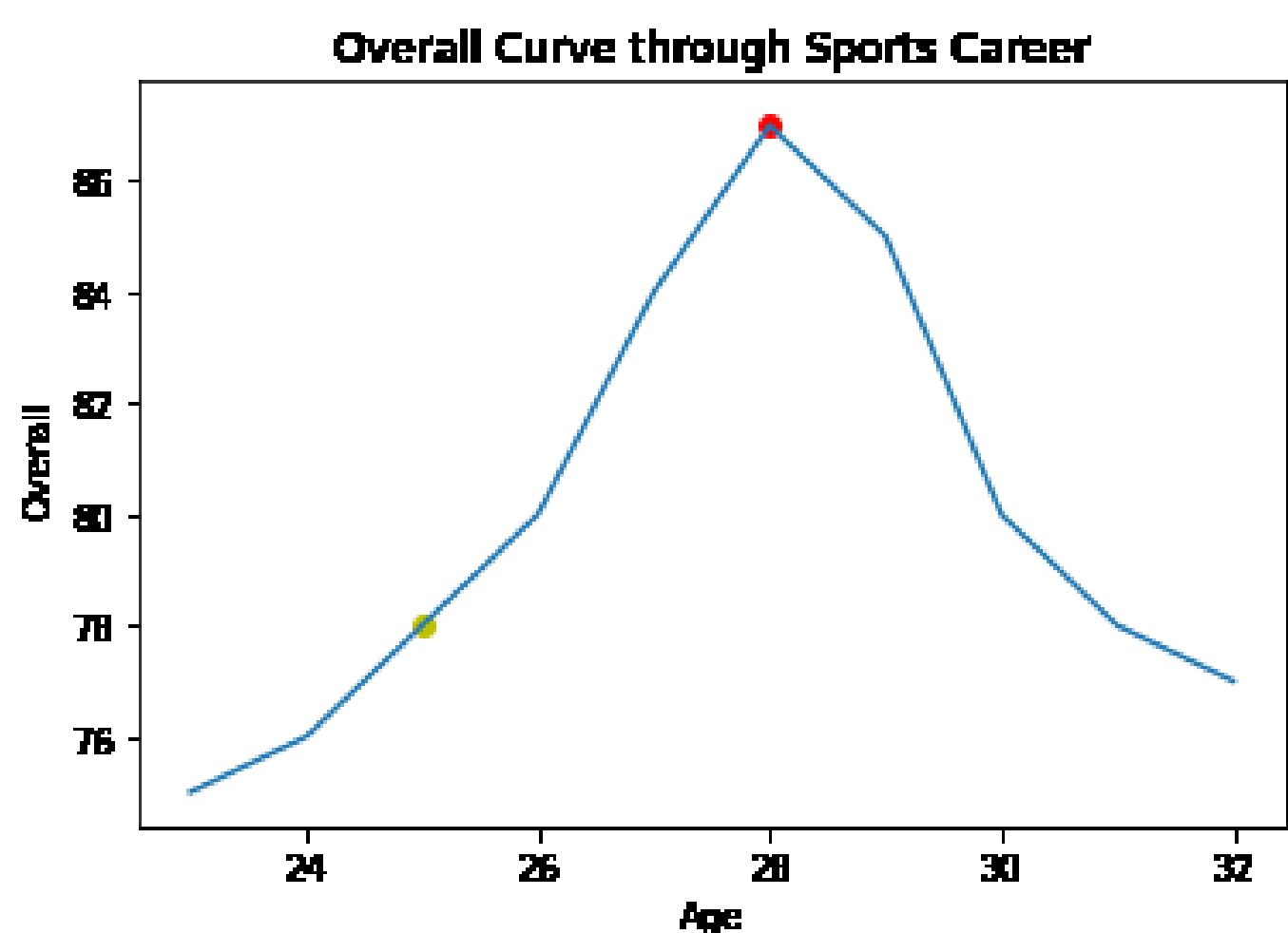
University of Chicago, Master of Computational Social Science Program

**MASTERS IN
COMPUTATIONAL
SOCIAL SCIENCE**
THE UNIVERSITY OF CHICAGO

Research Question

The primary purpose of my project is to predict soccer player’s potential. By using all numeric attributes which evaluate different aspects of a soccer player, I hope my model could predict the overall value of that player after three years.

To make it clear, the word ‘potential’ here means the peak of the overall value a player could ever reach in his sports career (overall value is a comprehensive index in the FIFA game to measure the general ability of a player).



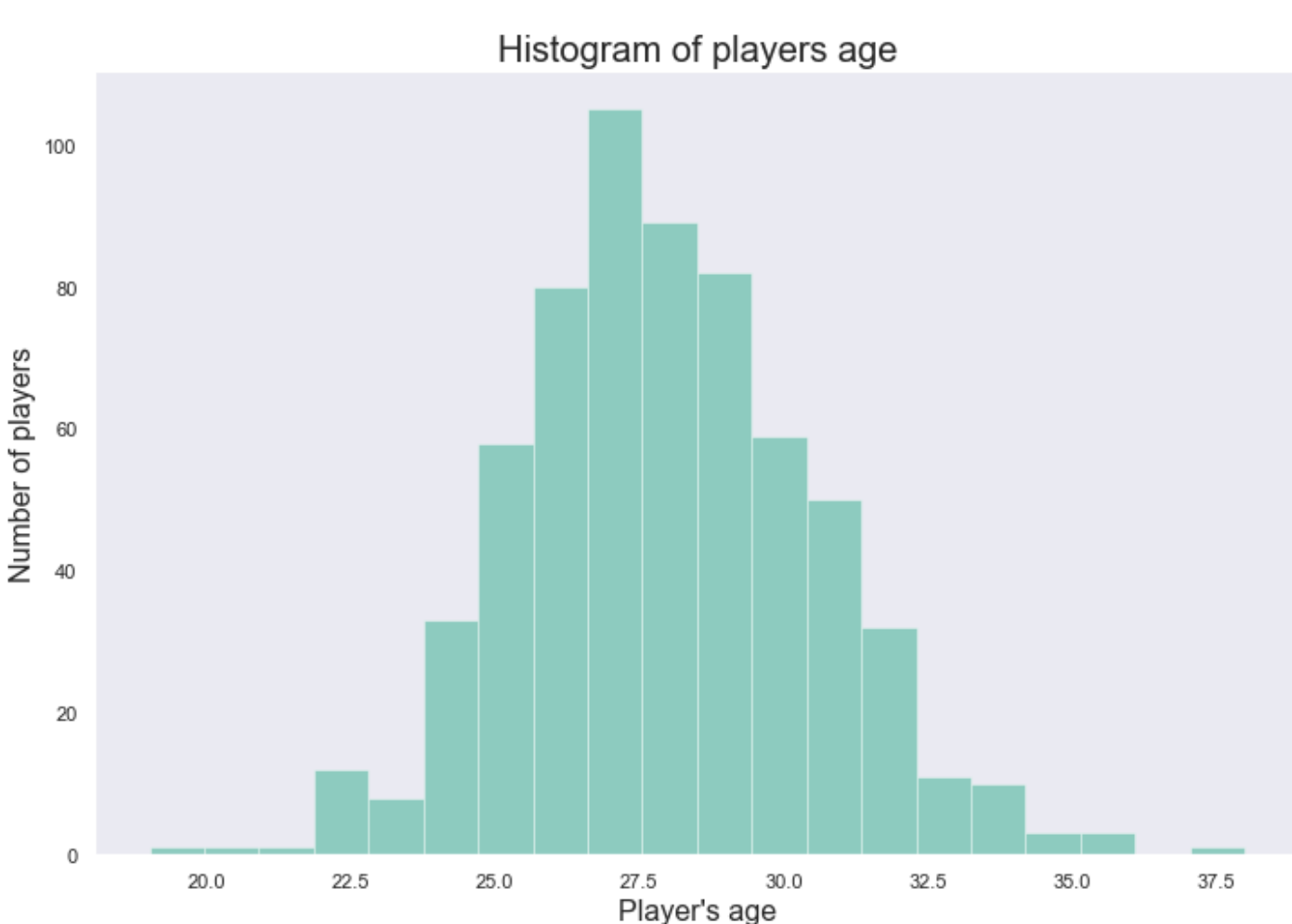
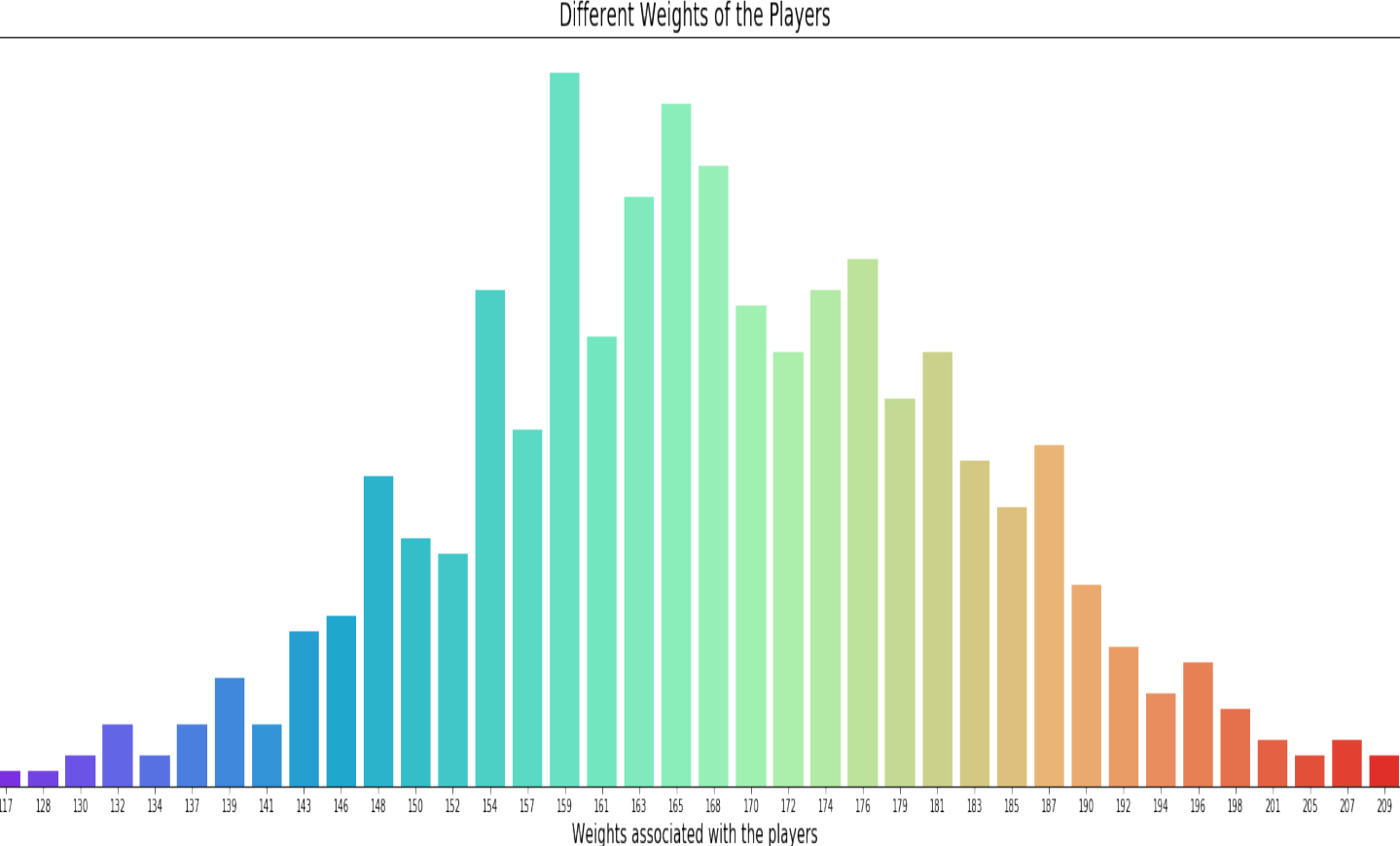
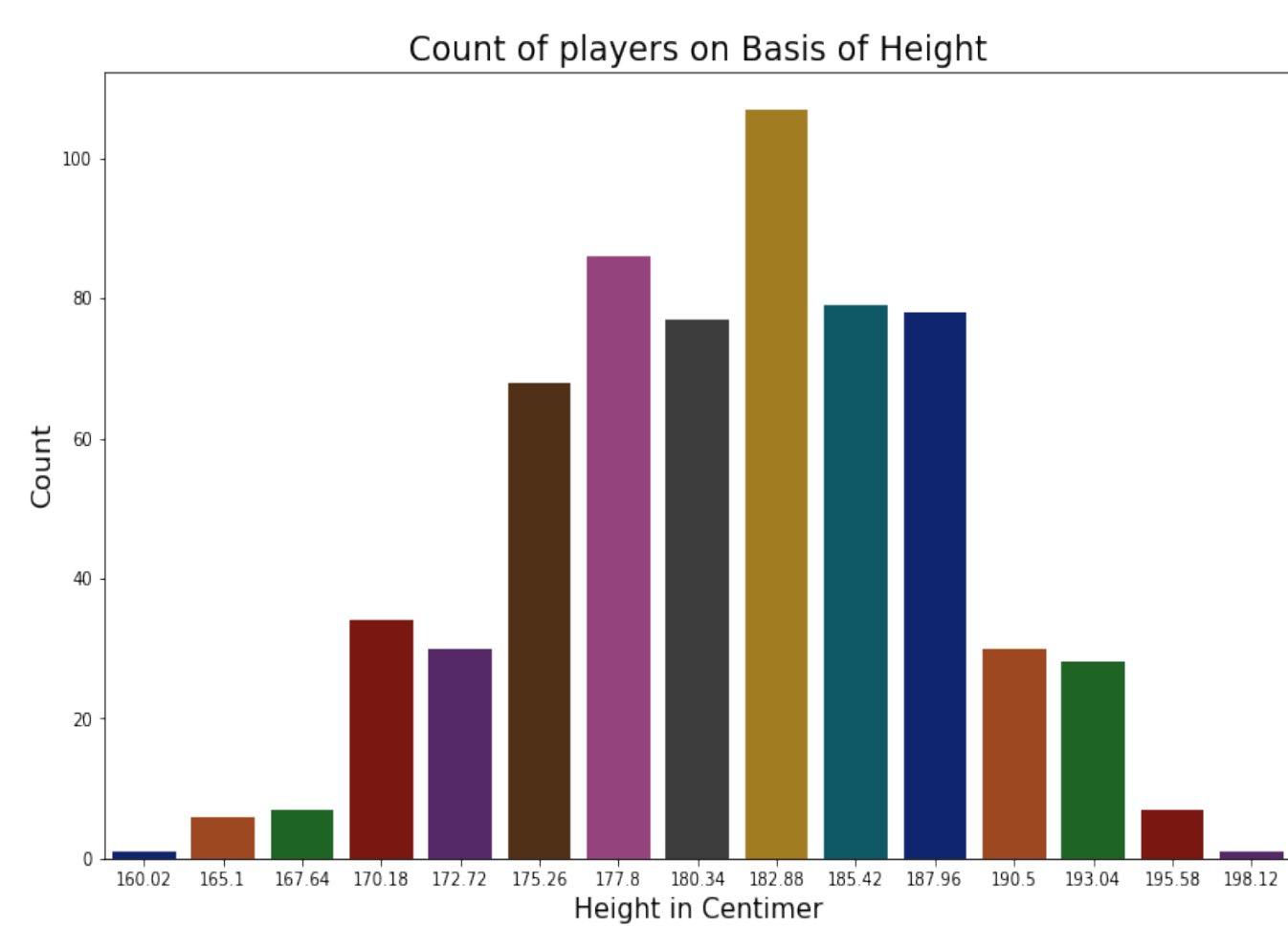
Data

Data Collection

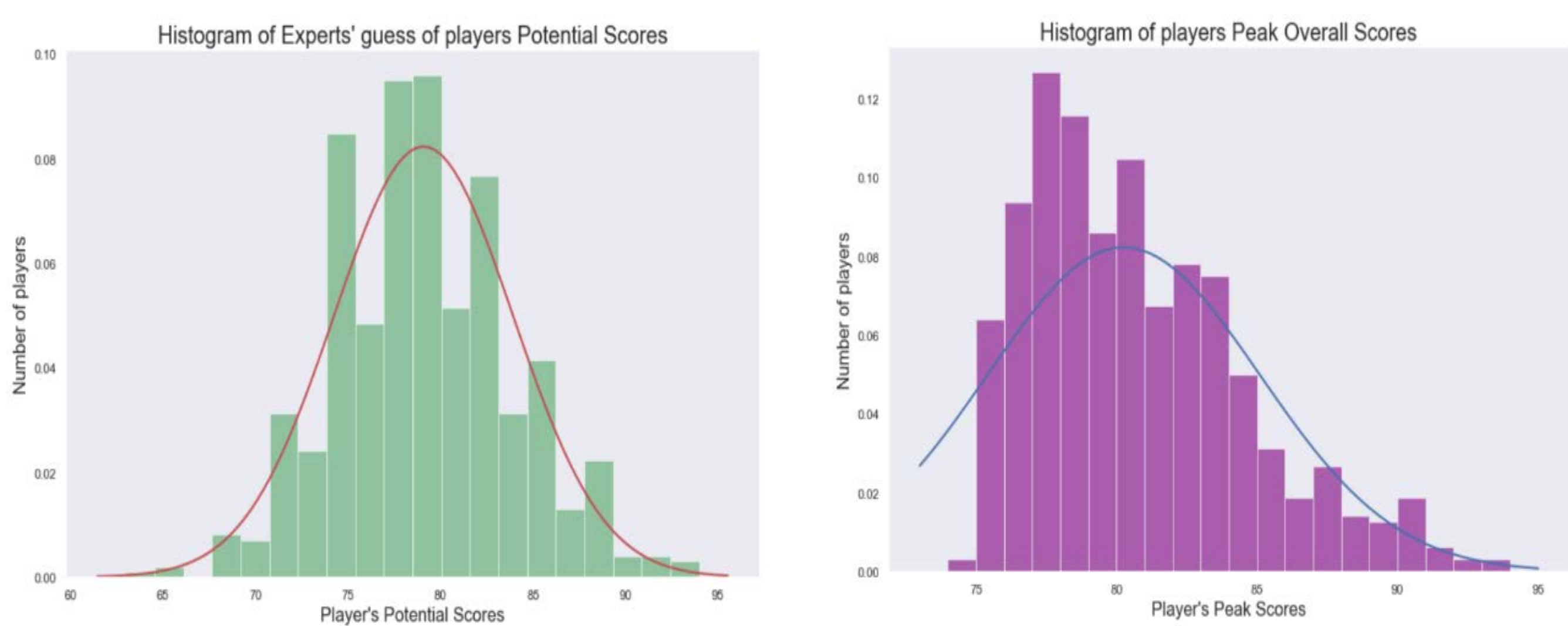
The data was scraped from the popular FIFA game website: www.sofifa.com. 639 players who appear from FIFA 10 to FIFA 19 constitute the final dataset.

Table 1. Summary Statistics

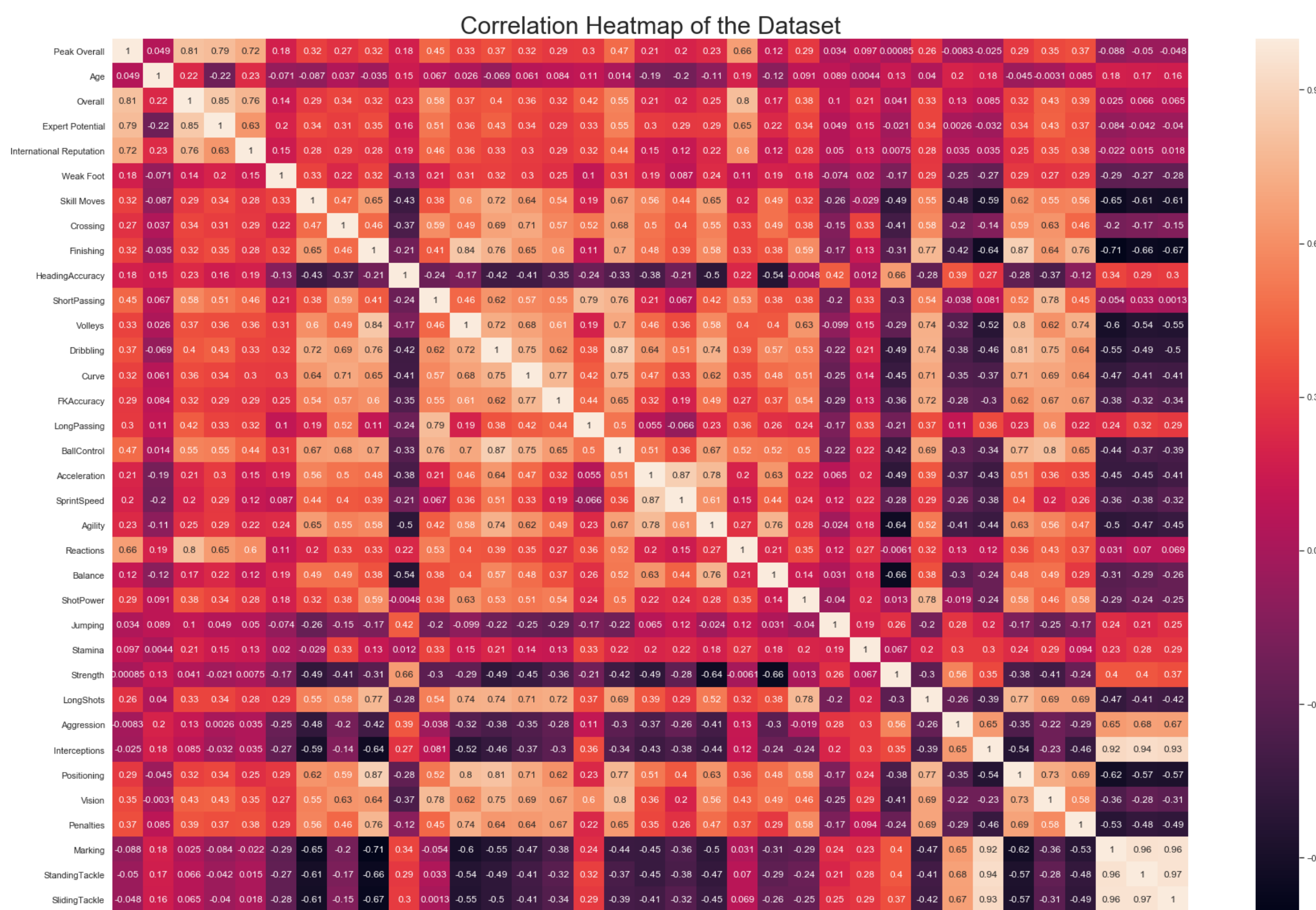
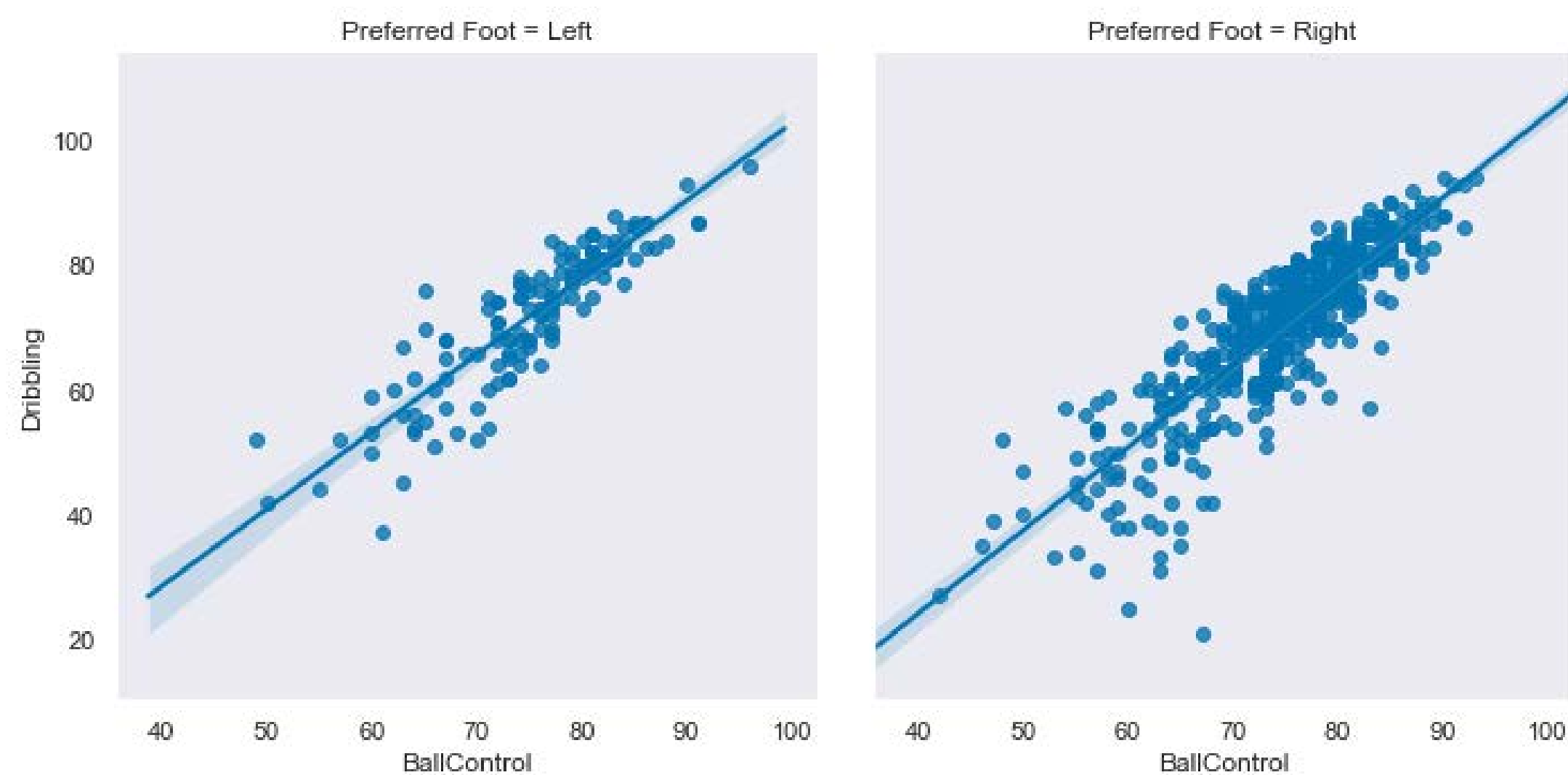
Category	Variable	Mean	Std	Min	Max
Basic	Peak Overall	80.26	3.89	74	94
	3 years before Age	24.91	2.70	16	35
	3 years before Overall	76.10	4.86	59	94
	Height(cm.)	181.4	6.45	160	198
	Weight(lbs.)	167.7	14.87	117	209
	Expert Potential	79.09	4.85	63	94
Attacking	International Reputation	1.67	0.85	1	5
	Weak Foot	3.20	0.69	1	5
	Skill Moves	2.82	0.77	2	5
	Crossing	64.63	13.89	16	90
	Finishing	58.93	17.44	10	95
	Heading Accuracy	65.53	12.12	24	95
Skill	Short Passing	73.21	7.73	48	92
	Volley	58.15	16.31	13	89
	Dribbling	70.46	12.45	21	96
	Curve	63.08	15.48	13	92
	FK Accuracy	58.17	16.05	10	91
	Long Passing	67.16	10.13	31	92
Movement	Ball Control	74.54	8.20	42	96
	Acceleration	73.90	10.12	34	95
	Sprint Speed	74.31	9.44	42	96
	Agility	71.86	12.06	30	95
	Reactions	74.16	6.82	54	92
	Balance	68.37	12.56	32	95
Power	Shot Power	71.52	10.52	24	94
	Jumping	70.74	10.28	33	95
	Stamina	75.96	8.20	51	94
Mentality	Strength	71.81	10.62	25	93
	Aggression	69.18	12.97	23	92
	Interceptions	60.17	20.10	14	91
	Positioning	64.56	16.15	12	94
	Vision	66.95	12.37	22	93
	Penalties	60.76	13.67	13	92
Defending	Marking	54.28	22.61	10	91
	Standing Tackle	59.30	22.03	11	91
	Sliding Tackle	55.98	22.78	11	92



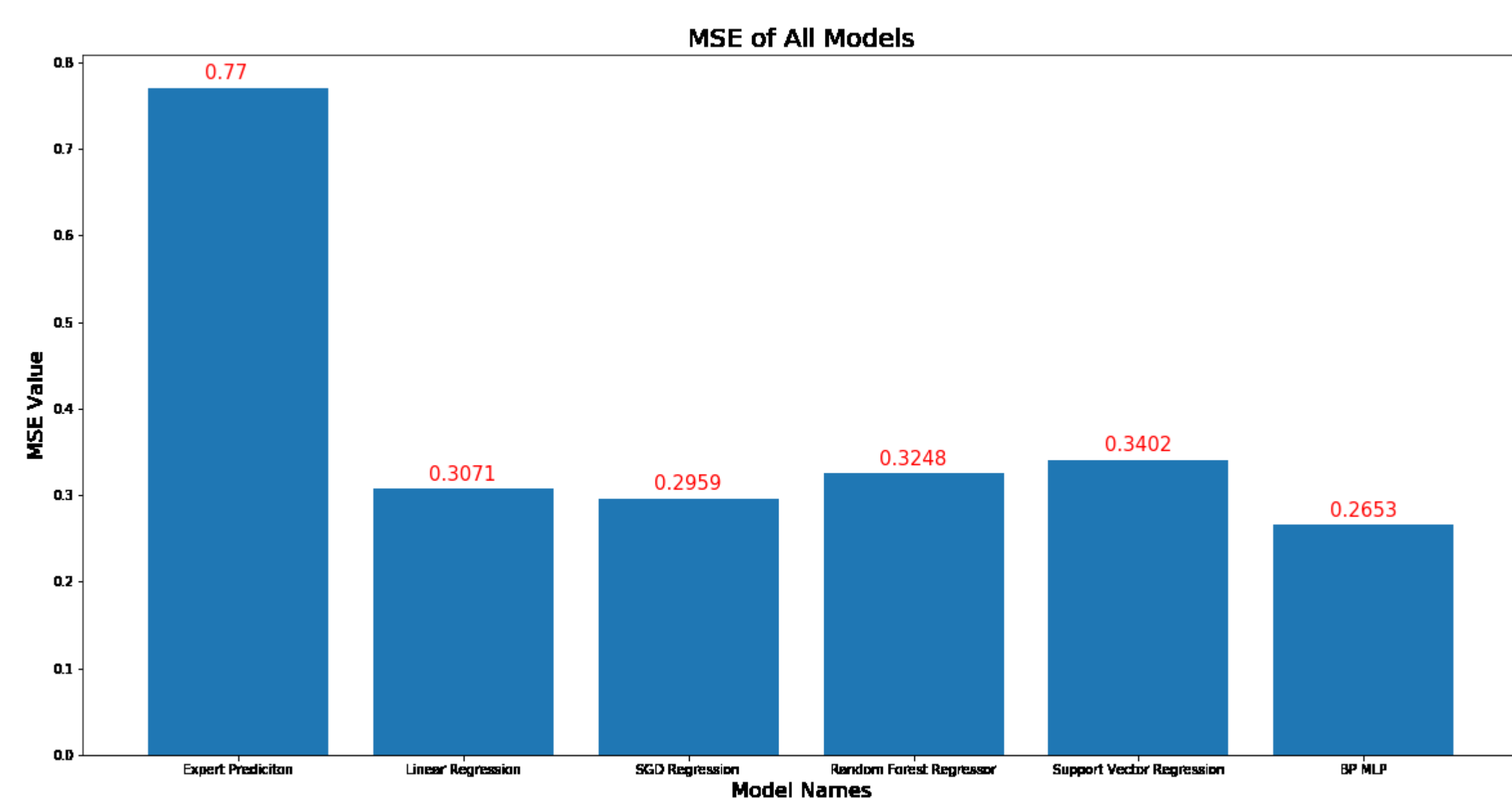
Justification for the Research Question



Variables Selection



Method



Results (Permutation Importance)



Table 2. Prediction Divergence

Underestimated Players							
Peak Overall	Age	Current Overall	Expert Potential	International Reputation	gap_1	gap_2	gap_3
76	26	62	63	1	0.10	-3.31	3.41
76	27	69	69	1	0.33	-1.78	2.12
76	24	65	69	1	0.21	-1.78	2.00
75	29	69	69	1	0.23	-1.53	1.76
79	25	68	70	1	-0.57	-2.29	1.73
75	25	69	71	1	0.67	-1.02	1.69
77	26	70	71	1	0.01	-1.53	1.54
77	25	69	71	1	-0.02	-1.53	1.51
76	27	71	71	1	0.20	-1.27	1.48
78	26	74	74	2	0.45	-1.02	1.47

Table 3. Predicted Young Talents

Predictions Comparison (2019)				
Rank	Current Overall	Expert Potential	MLP Potential	Most Undervalued
1	H. Kane	K. Mbappé	H. Kane	J. Damm
2	P. Dybala	P. Dybala	P. Pogba	F. Boulaya
3	K. Mbappé	M. de Ligt	K. Mbappé	L. Krejčí
4	P. Pogba	M. Škriniar	P. Dybala	S. Skrzybski
5	S. Umtiti	H. Kane	H. Son	T. Goiginger
6	R. Sterling	S. Umtiti	J. Kimmich	Hwang Ui Jo
7	M. Icardi	O. Dembélé	R. Lukaku	O. Hernández
8	H. Son	L. Sané	D. Alli	P. Ntep
9	A. Laporte	Vinicius Júnior	R. Sterling	S. Kittel
10	M. Verratti	P. Pogba	M. Icardi	Carles Gil

Summary

Conclusion:

MLP model performs the best.

Current overall value, international reputation, Reaction, Age, Volley Ability, Ball Control, and Long Shot Ability have a significant impact on the potential value.

Practical Significance:

For role players, MLP model’s prediction is much more accurate than experts’.