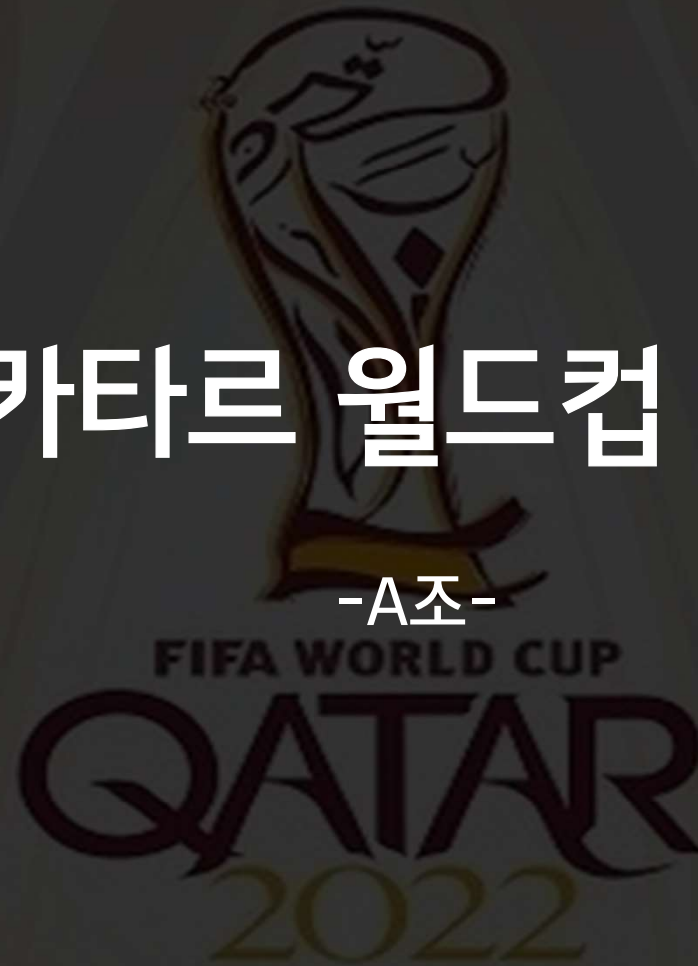


# 2022 카타르 월드컵 순위 예측

-A조-



QATAR 2022

# 목차

프로젝트 개요

데이터 분석

데이터 예측

보완점

Q/A



QATAR 2022

# 프로젝트 개요 | 팀 소개



KANG



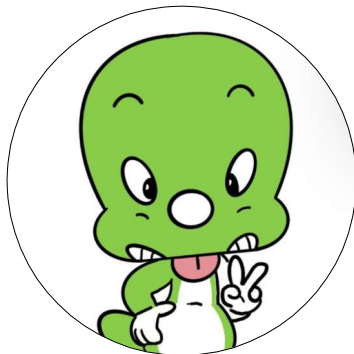
KIM



LEE



HWANG



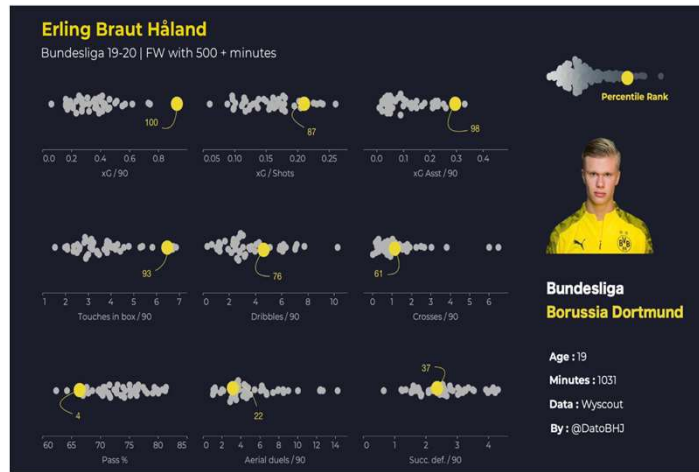
HAN

데이터 수집					크롤링 PL 팀별 베스트 선수 선발 (15명)
데이터 전처리	월드컵 32개국 선수	FIFA 22 데이터 전처리	FIFA 22 데이터 전처리	FIFA 22 데이터 전처리	
데이터 분석	연봉과 승리 및 득점 연관성 분석		슈팅수 등 4개 항목과 승리 및 득점 연관성 분석	포지션 별 능력치와 승리 및 득점 연관성 분석	슈팅수 등 4개 항목과 승리 및 득점 연관성 분석
데이터 예측		알고리즘			

# 프로젝트 개요 | 배경 및 목적



21'-22' PL 득점왕 손흥민 (23골)



축구는 데이터 싸움이다

xG: 기대 득점  
xA: 기대 어시스트  
팀 점유율 등..



2022년 11월 카타르 월드컵 개막

# 데이터 분석 | 데이터 수집

---

[ 선수 실력 ] [ 경기 영향 요인 ] [ 테스트 데이터셋 ] [ 예측 데이터셋 ]

# 데이터 분석 | 데이터 수집

[ 선수 실력 ] [ 경기 영향 요인 ] [ 테스트 데이터셋 ] [ 예측 데이터셋 ]



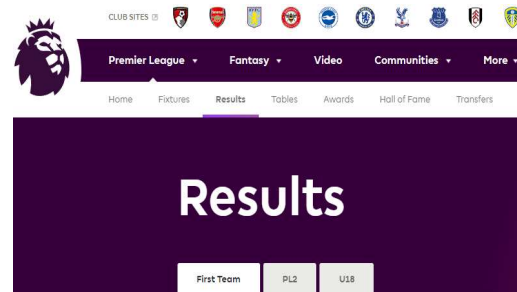
약 20,000명의 축구선수

슈팅 능력, 드리블 능력...  
등 35개 항목



21-22 PL 272 경기 결과

경기별 슈팅, 유효슈팅...  
등 8개 항목



21-22 PL 272 경기 결과



2022 월드컵 대진표



# 데이터 분석 | 데이터 전처리

## [ 선수 실력 ]



약 20,000명의 축구선수

슈팅 능력, 드리블 능력...

등 35개 항목

players\_22.csv

검색(Alt+Q)

lee sueng jun

파일 홈 삽입 레이아웃 수식 데이터 검토 보기 도움말

자동 줄 바꿈

표준 나쁨 보통 좋음 경고문

조건부 서식

스타일

삽입 삭제 서식

자동 합계

재우기

정렬 및 찾기 및 필터

선택

문서 보기

sofifa\_id

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
sofifa_id	player_url	short_name	long_name	player_pos	overall	potential	value_eur	wage_eur	age	dob	height_cm	weight_kg	club_team	club_name	league_name	league_level	club_posit	club_jersey	club_loane	club_join	club_contr
158023	https://sofifa.com/players/158023/	L. Messi	Lionel AncRW	ST	CF	93	93	78000000	320000	34	170	72	73	Paris Saint-Germain	French Lig	1	RW	30			2023
188545	https://sofifa.com/players/188545/	R. Lewandowski	Robert Lewandowski	ST	ST	92	92	1.2E+08	270000	32	185	81	21	FC Bayern Munich	German 1.	1	ST	9			2023
20801	https://sofifa.com/players/20801/	Cristiano Ronaldo	Cristiano Ronaldo	ST	LW	91	91	45000000	270000	36	187	83	11	Manchester United	English Pr	1	ST	7			2023
190871	https://sofifa.com/players/190871/	Neymar Jr	Neymar Jr	LW	CAM	91	91	1.29E+08	270000	29	175	68	73	Paris Saint-Germain	French Lig	1	LW	10			2025
192985	https://sofifa.com/players/192985/	Kevin De Bruyne	Kevin De Bruyne	CM	CAM	91	91	1.26E+08	350000	30	181	70	10	Manchester City	English Pr	1	RCM	17			2025
200389	https://sofifa.com/players/200389/	J. Oblak	Jan Oblak	GK	GK	91	93	1.12E+08	130000	28	188	87	240	Atletico de Madrid	Spain Prin	1	GK	13			2023
231747	https://sofifa.com/players/231747/	K. Mbappe	Kylian Mbappe	ST	LW	91	95	1.94E+08	230000	22	182	73	73	Paris Saint-Germain	French Lig	1	ST	7			2022
167495	https://sofifa.com/players/167495/	M. Neuer	Manuel Neuer	GK	GK	90	90	13500000	86000	35	193	93	21	FC Bayern Munich	German 1.	1	GK	1			2023
192448	https://sofifa.com/players/192448/	M. ter Stegen	Marc-André ter Stegen	GK	GK	90	92	99000000	250000	29	187	85	241	FC Barcelona	Spain Prin	1	GK	1			2025
202126	https://sofifa.com/players/202126/	H. Kane	Harry Kane	ST	ST	90	90	1.3E+08	240000	27	188	89	18	Tottenham Hotspur	English Pr	1	ST	10			2024
215914	https://sofifa.com/players/215914/	N. Kanté	N'Golo Kanté	CM	CM	90	90	1E+08	230000	30	168	70	5	Chelsea	English Pr	1	RCM	7			2023
165153	https://sofifa.com/players/165153/	K. Benzema	Karim Benzema	CF	ST	89	89	66000000	350000	33	185	81	243	Real Madrid	Spain Prin	1	CF	9			2023
192119	https://sofifa.com/players/192119/	T. Courtois	Thibaut Courtois	GK	GK	89	91	85500000	250000	29	199	96	243	Real Madrid	Spain Prin	1	GK	1			2026
200104	https://sofifa.com/players/200104/	H. Son	Heung-min Son	LM	CF, LV	89	89	1.04E+08	220000	28	183	78	18	Tottenham Hotspur	English Pr	1	LW	7			2025
200145	https://sofifa.com/players/200145/	C. Casemiro	Carlos Casemiro	CDM	CDM	89	89	88000000	310000	29	185	84	243	Real Madrid	Spain Prin	1	CDM	14			2025
203376	https://sofifa.com/players/203376/	V. van Dijk	Virgil van Dijk	CB	CB	89	89	86000000	230000	29	193	92	9	Liverpool	English Pr	1	LCB	4			2025
208722	https://sofifa.com/players/208722/	S. Mané	Sadio Mané	LW	LW	89	89	1.01E+08	270000	29	175	69	9	Liverpool	English Pr	1	LW	10			2023
209331	https://sofifa.com/players/209331/	M. Salah	Mohamed Salah	RW	RW	89	89	1.01E+08	270000	29	175	71	9	Liverpool	English Pr	1	RW	11			2023
210257	https://sofifa.com/players/210257/	E. Ederson	Ederson	GK	GK	89	91	94000000	200000	27	188	86	10	Manchester City	English Pr	1	GK	31			2026
212622	https://sofifa.com/players/212622/	J. Kimmich	Joshua Kimmich	CDM	RB	89	90	1.08E+08	160000	26	177	75	21	FC Bayern Munich	German 1.	1	RDM	6			2025
212831	https://sofifa.com/players/212831/	A. Alisson	Alisson	Ra	GK	89	90	82000000	190000	28	191	91	9	Liverpool	English Pr	1	GK	1			2027
230621	https://sofifa.com/players/230621/	G. Donnarumma	Gianluigi Donnarumma	GK	GK	89	93	1.2E+08	110000	22	196	90	73	Paris Saint-Germain	French Lig	1	GK	50			2026
155862	https://sofifa.com/players/155862/	S. R. Sergio	Rafael Sergio	CB	CB	88	88	24000000	115000	35	184	82	73	Paris Saint-Germain	French Lig	1	LCB	4			2023
176580	https://sofifa.com/players/176580/	L. Suñez	Luis Suñez	ST	ST	88	88	44500000	135000	34	182	83	240	Atletico de Madrid	Spain Prin	1	RS	9			2022
182521	https://sofifa.com/players/182521/	T. Kroos	Toni Kroos	CM	CM	88	88	75000000	310000	31	183	76	243	Real Madrid	Spain Prin	1	LCM	8			2023
192505	https://sofifa.com/players/192505/	R. Lukaku	Romelu Lukaku	ST	ST	88	88	93500000	260000	28	191	94	5	Chelsea	English Pr	1	ST	9			2026
193041	https://sofifa.com/players/193041/	K. Navas	Keylor Navas	GK	GK	88	88	15500000	130000	34	185	80	73	Paris Saint-Germain	French Lig	1	SUB	1			2024
202652	https://sofifa.com/players/202652/	R. Sterling	Raheem Sterling	LW	RW	88	89	1.03E+08	290000	26	170	69	11	Manchester City	English Pr	1	SUB	7			2023
212198	https://sofifa.com/players/212198/	B. Fernandes	Bruno Fernandes	CM	CM	88	89	1.08E+08	250000	26	179	69	11	Manchester United	English Pr	1	CAM	18			2025
239085	https://sofifa.com/players/239085/	E. Haaland	Erling Haaland	ST	ST	88	93	1.38E+08	110000	20	194	94	22	Borussia Dortmund	German 1.	1	RS	9			2024
153079	https://sofifa.com/players/153079/	S. Agüero	Sergio Agüero	ST	ST	97	87	31000000	260000	33	173	70	241	FC Barcelona	Spain Prin	1	ST	19			2023

# 데이터 분석 | 데이터 전처리

## [ 선수 실력 ]

name	overall	weak_Foot	skill_moves	pace	shooting	passing	dribbling	defending	physic	attacking_crossing	attacking_finishing	...
L. Messi	93	4	4	85	92	91	95	34	65	85	95	...

선수 포지션별

슈팅 능력, 드리블 능력...  
등 35개 항목

Using Virgil van Dijk's FIFA 20 in-game scores, we can use the Positional Coefficients for centre-backs to see exactly how his overall rating of 90 was calculated:

Statistic	CB Coefficient	Van Dijk	CB Overall
Defensive Awareness	0.15	91	13.65
Standing Tackle	0.15	92	13.8
Sliding Tackle	0.15	86	12.9
Heading Accuracy	0.1	86	8.6
Strength	0.1	92	9.2
Aggression	0.08	82	6.56
Interceptions	0.08	90	7.2
Short Passing	0.05	79	3.95
Ball Control	0.05	76	3.8
Reactions	0.05	88	4.4
Jumping	0.04	90	3.6
TOTAL	1	952	87.66

→ 세가지 항목 간의 비교를 통한 가중치 근거 설정

	1	2	3	4
1	Team	DF	MF	FW
2	Arsenal	75.714	76.875	79.6
3	Aston Villa	78.28571	74.57143	80.33333
4	Brentford	72.1667	74	72.6667
5	Brighton & Hove Albion	76	75	75.66667
6	Chelsea	80.57143	83.25	83
7	Crystal Palace	74	75.5	76.4
8	Everton	78.125	78.5	76.5
9	Leeds United	76.1667	75.14286	75
10	Leicester City	78.14286	79	79
11	Liverpool	83	81	83.3333



# 데이터 분석 | 데이터 전처리

## [ 경기 영향 요인 ]



## 21-22 PL 272 경기 결과

경기별 슈팅, 유효슈팅...

등 8개 항목

[illegible]

# 데이터 분석 | 데이터 전처리

[ 선수 실력 ] [ 경기 영향 요인 ] [ 테스트 데이터셋 ] [ 예측 데이터셋 ]



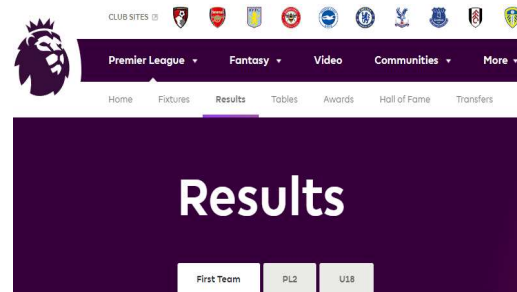
약 631명의 축구선수

- 포지션별 능력 평균치



21-22 PL 272 경기 결과

- 슈팅 수
- 유효 슈팅 수
- 코너킥
- 파울
- 연봉



21-22 PL 272 경기 결과



2022 월드컵 대진표

# 데이터 분석 | 상관 분석 - 회귀 분석

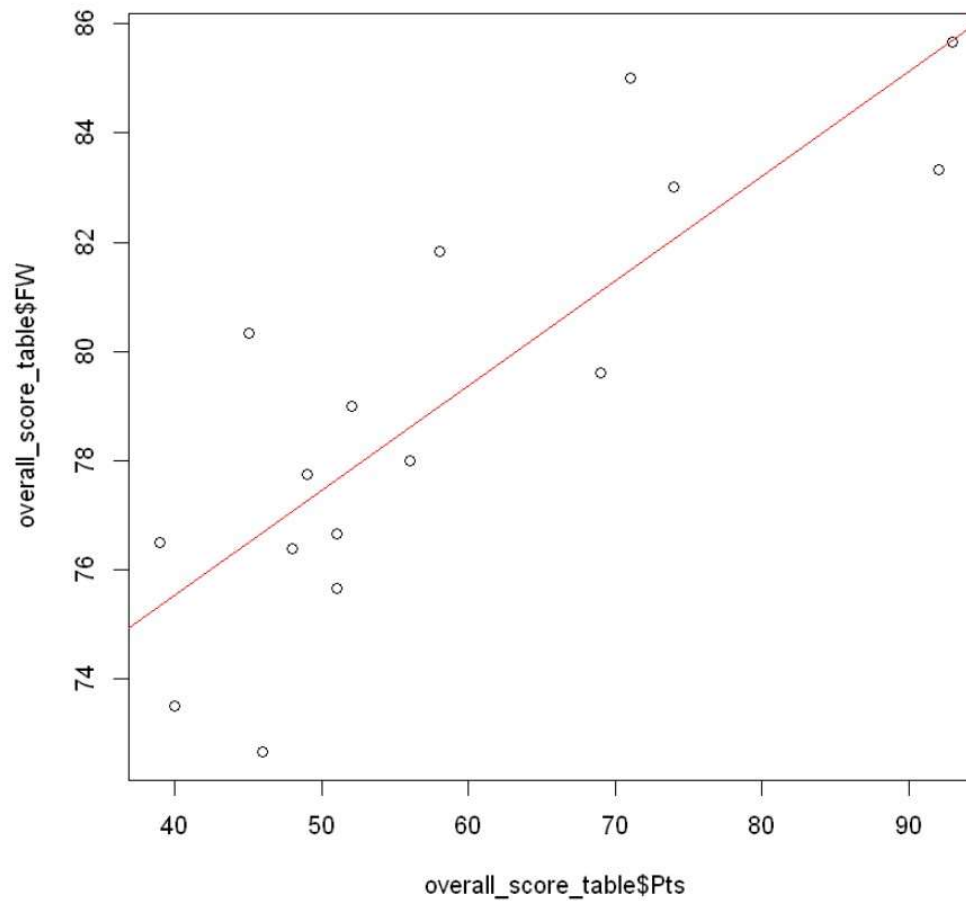
## [ 선수 실력 ]



약 631명의 축구선수

- 포지션별 능력 평균치

FW



# 데이터 분석 | 상관 분석 - 회귀 분석

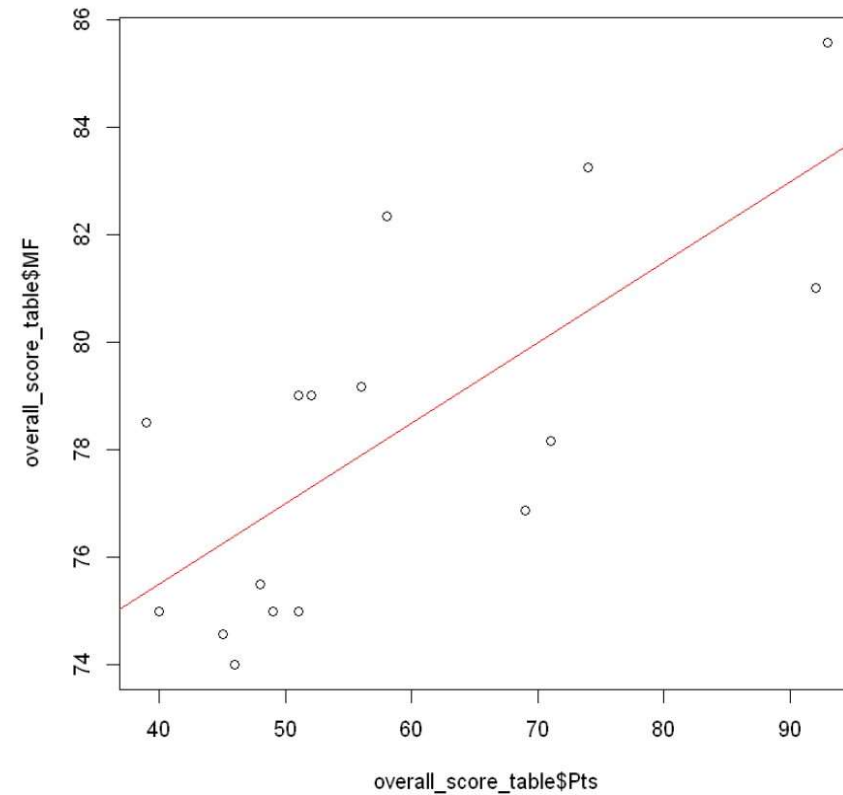
## [ 선수 실력 ]



약 631명의 축구선수

- 포지션별 능력 평균치

MF



# 데이터 분석 | 상관 분석 - 회귀 분석

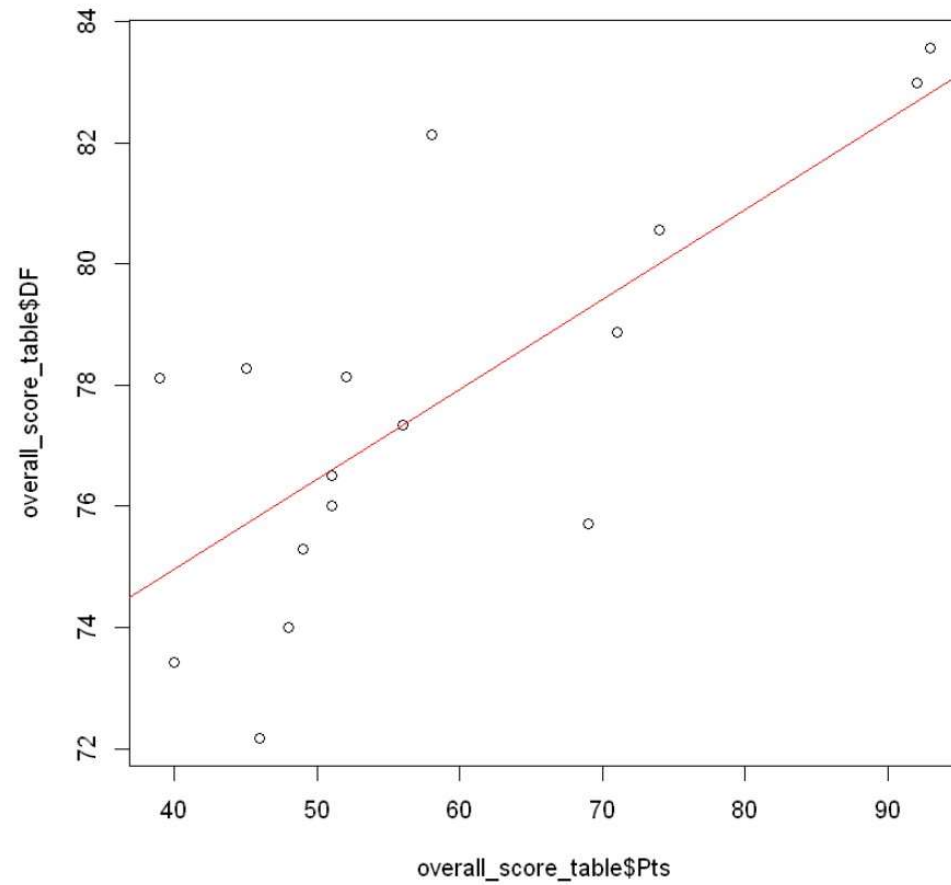
## [ 선수 실력 ]



약 631명의 축구선수

- 포지션별 능력 평균치

DF



# 데이터 분석 | 상관 분석 - 회귀 분석

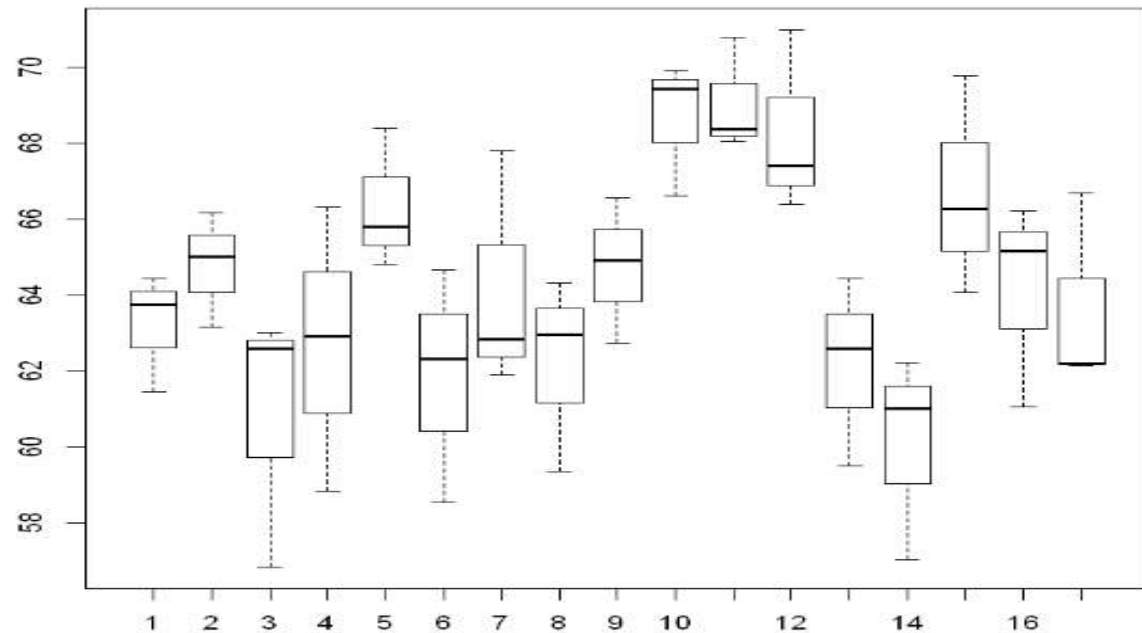
## [ 선수 실력 ]



약 631명의 축구선수

- 포지션별 능력 평균치

- 경기 승리 연관성 분석



Aston\_Villa,Leeds\_Unite,Brighton\_and\_Hove\_Albian,Chelsea,Crystal\_Palace,Everton,,Leicester\_City,Brentford,Tottenham\_Hotspur,Liverpool,Manchester\_City,Manchester\_United,Newcastle\_United,Southampton,,Arsenal,West\_Ham\_United,Wolverhampton\_Wanderers|

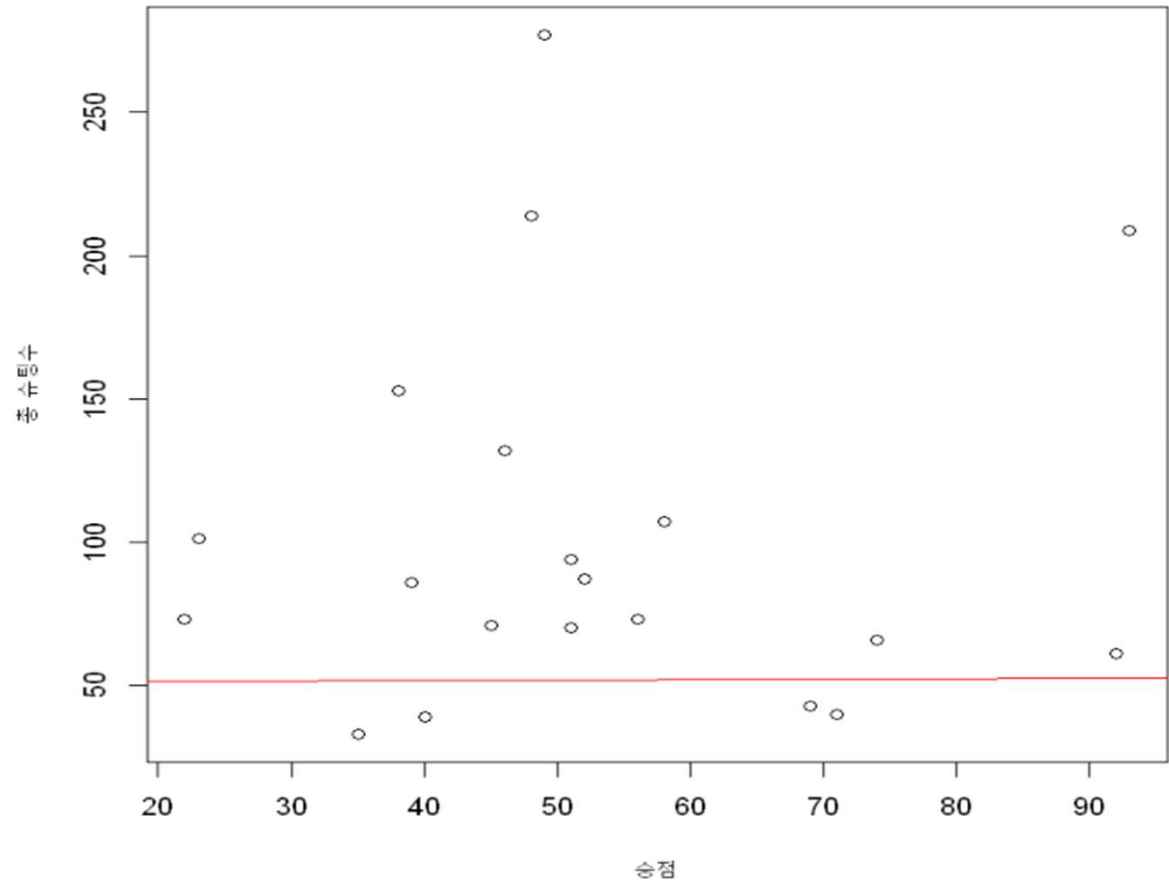
# 데이터 분석 | 상관 분석 - 회귀 분석

## [ 경기 영향 요인 ]



21-22 PL 272 경기 결과

- 슈팅 수
- 유효 슈팅 수
- 코너킥
- 파울
- 연봉





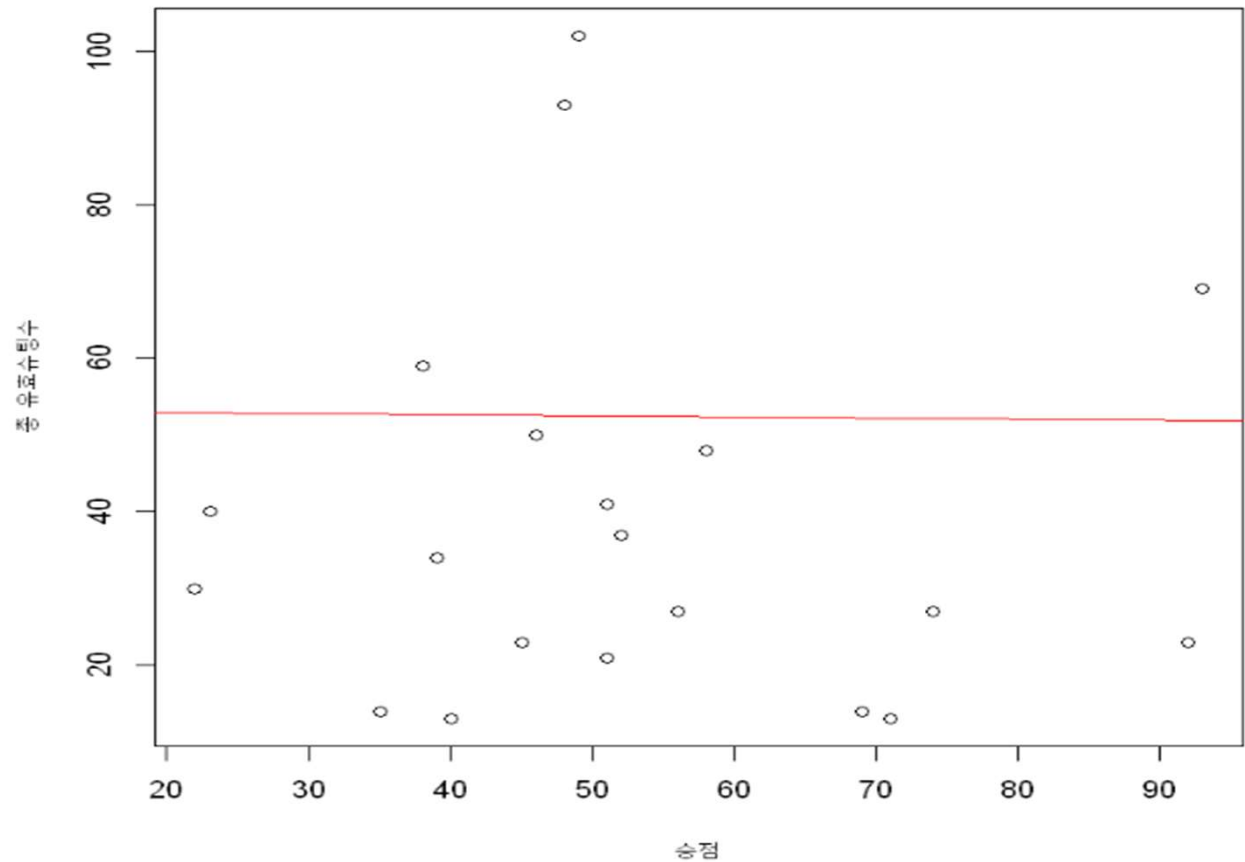
# 데이터 분석 | 상관 분석 - 회귀 분석

## [ 경기 영향 요인 ]



21-22 PL 272 경기 결과

- 슈팅 수
- 유효 슈팅 수
- 코너킥
- 파울
- 연봉



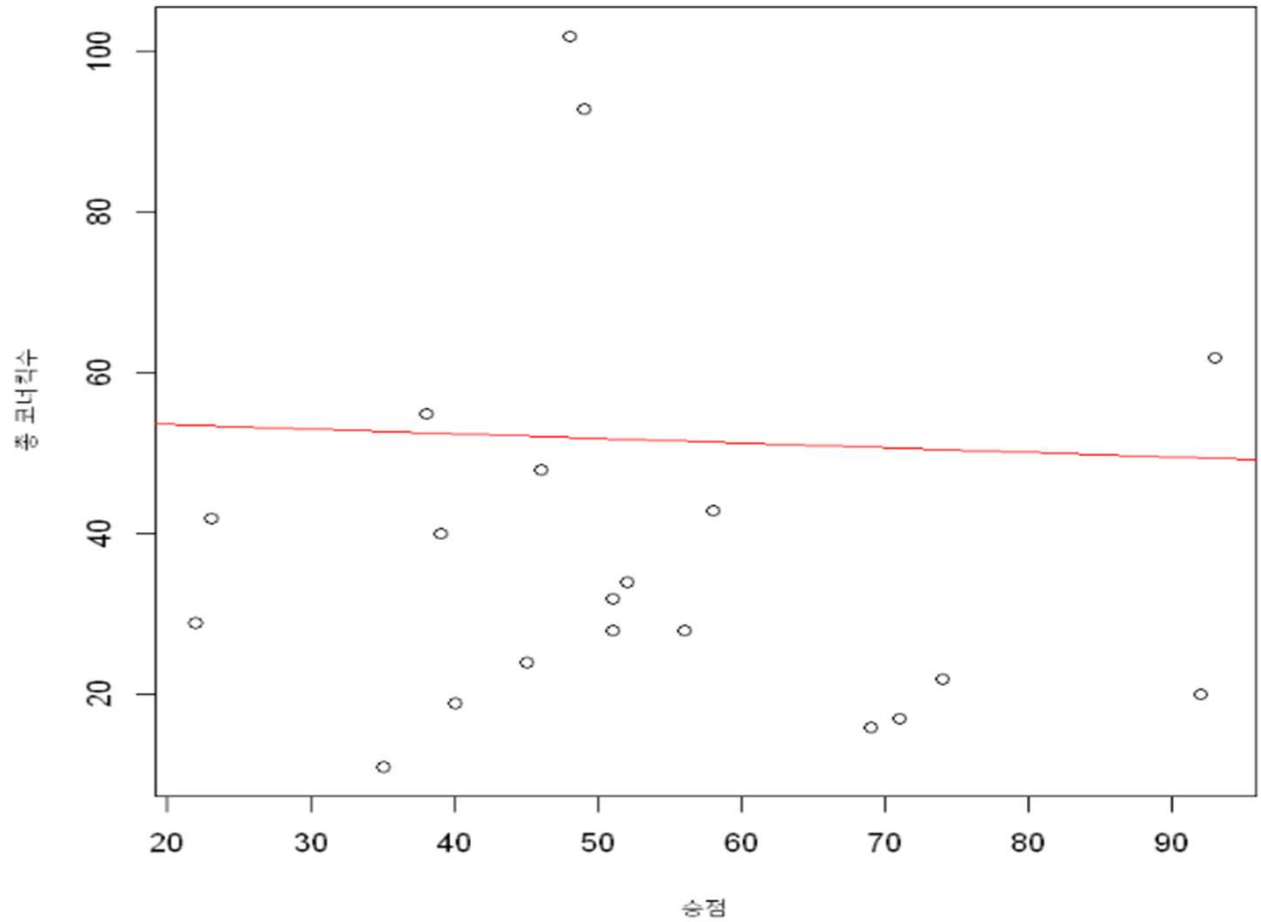
# 데이터 분석 | 상관 분석 - 회귀 분석

## [ 경기 영향 요인 ]



21-22 PL 272 경기 결과

- 슈팅 수
- 유효 슈팅 수
- 코너킥
- 파울
- 연봉



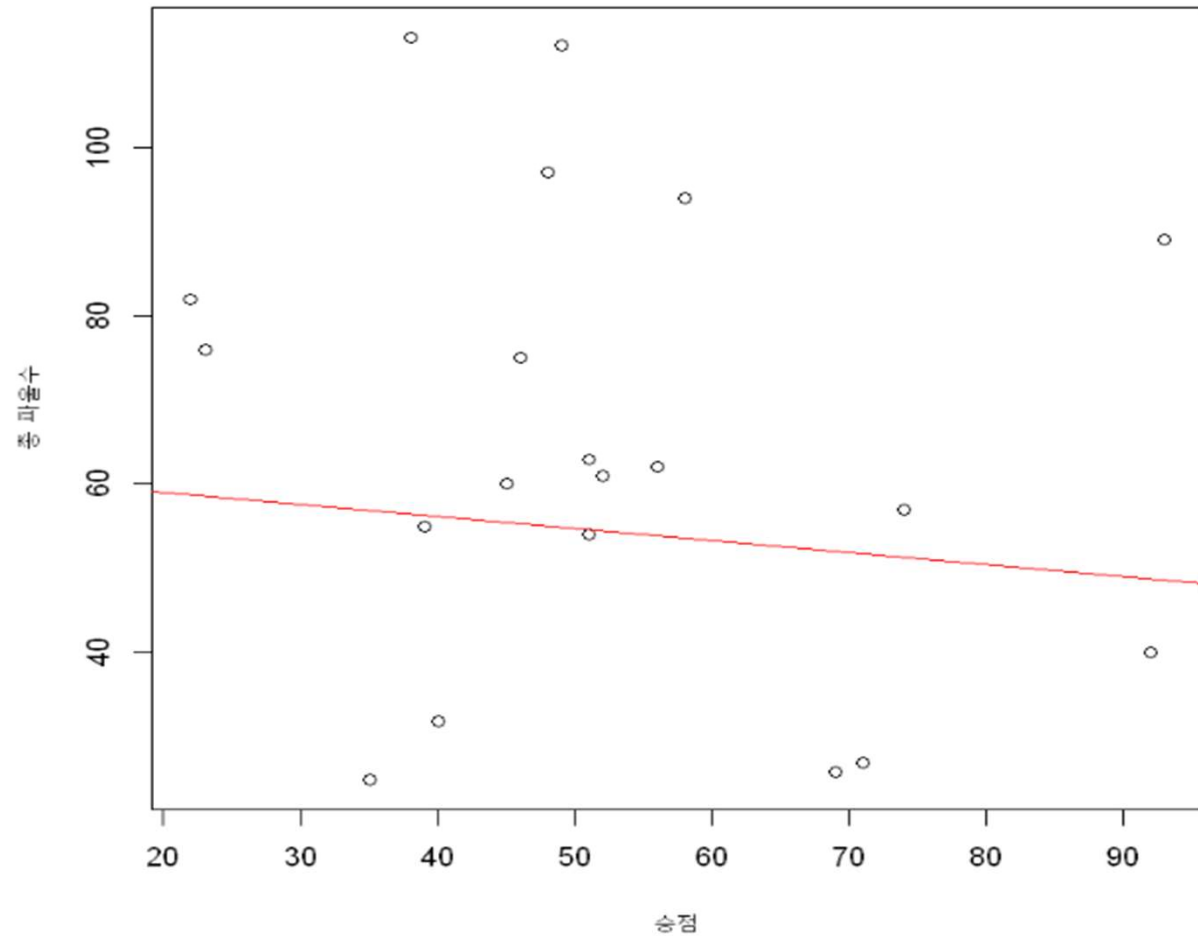
# 데이터 분석 | 상관 분석 - 회귀 분석

## [ 경기 영향 요인 ]



21-22 PL 272 경기 결과

- 슈팅 수
- 유효 슈팅 수
- 코너킥
- 파울
- 연봉



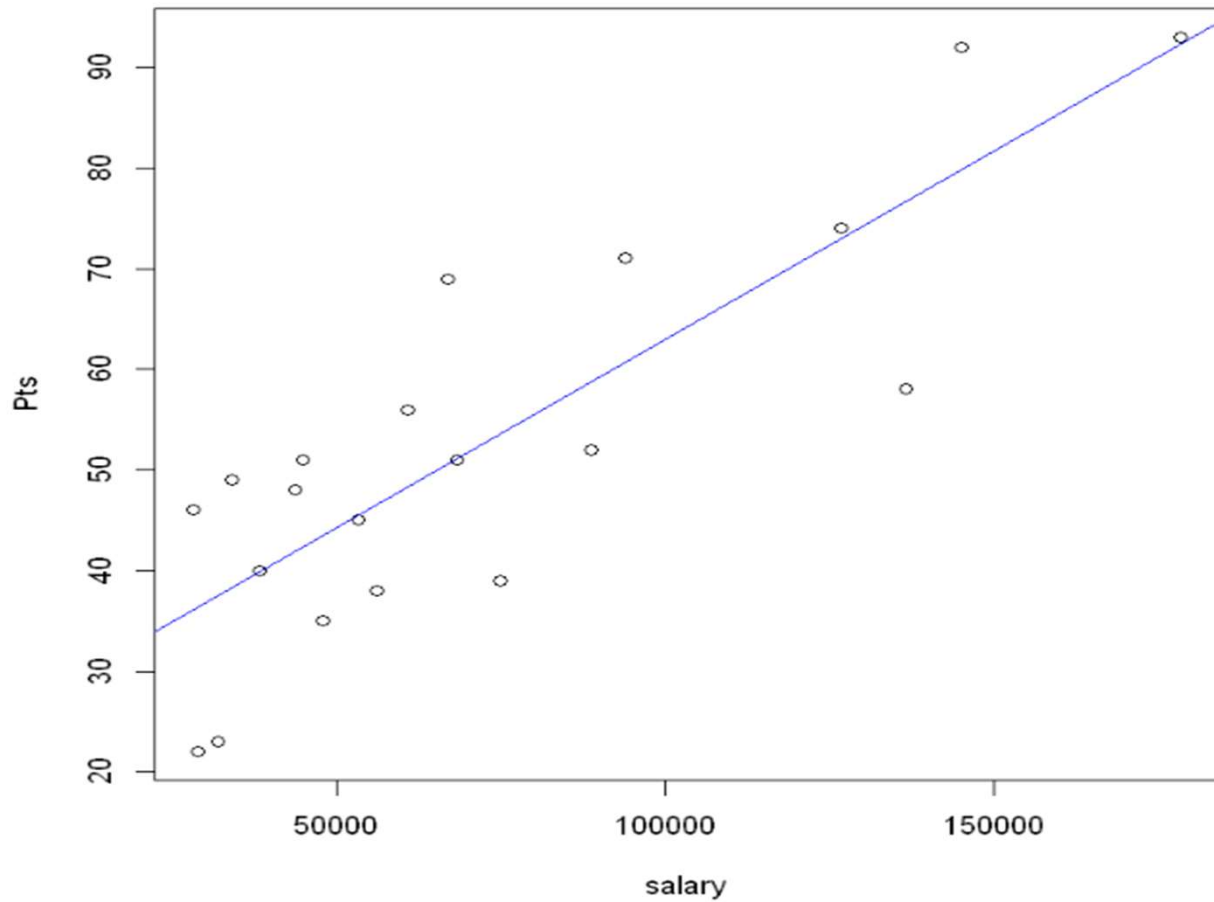
# 데이터 분석 | 상관 분석 - 회귀 분석

## [ 경기 영향 요인 ]



21-22 PL 272 경기 결과

- 슈팅 수
- 유효 슈팅 수
- 코너킥
- 파울
- 연봉



# 데이터 분석 | 분석 결과

[ 선수 실력 ] [ 경기 영향 요인 ] [ 테스트 데이터셋 ] [ 예측 데이터셋 ]



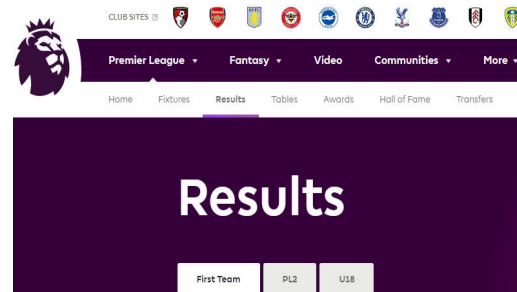
약 631명의 축구선수

- 포지션별 능력 평균치



21-22 PL 272 경기 결과

- 슈팅 수
- 유효 슈팅 수
- 코너킥
- 파울
- 연봉



21-22 PL 272 경기 결과



2022 월드컵 대진표

# 데이터 분석 | 분석 결과

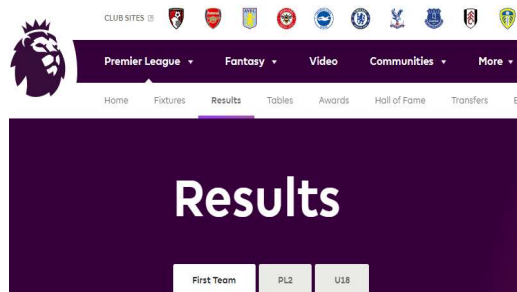
## [ 선수 실력 ]



약 631명의 축구선수

- **포지션별 능력 평균치**

## [ 테스트 데이터셋 ]



21-22 PL 272 경기 결과

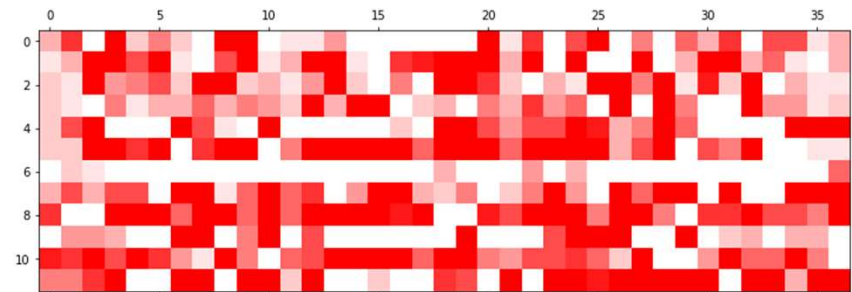
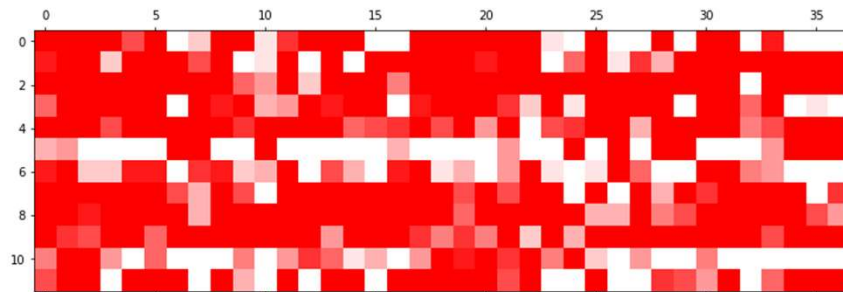
## [ 예측 데이터셋 ]



2022 월드컵 대진표

# 데이터 예측 | 예측 모델

[ 승리한 경기와 패배한 경기에 대해 각각 가중치와 능력치 차이를 곱한 결과 ]



공격수와 수비수의 능력치 차이가 클수록 승리할 확률이 올라가는 것을 확인할 수 있다.







# 데이터 예측 | 예측 모델

최종 ML 모델 정확도 : 52%

			실제 결과	예측결과		
				Win	Draw	Lose
8위 vs 16위			2:1	57.5%	13.7%	28.8%
11위 vs 5위			2:0	6.5%	11.5%	82.0%
1위 vs 14위			3:2	90.5%	3.5%	5.9%

# 데이터 예측 | 예측 결과



	대한민국 예측결과		
	Win	Draw	Lose
	40.70%	22.60%	36.70%
	52.70%	17.00%	30.30%
	39.20%	25.00%	42.10%
	6.30%	7.40%	86.20%

# 데이터 예측 | 예측 결과

GROUPS		ROUND OF 16		QUARTER-FINALS		SEMI-FINALS		FINAL	
<div> <div>A</div> <div>B</div> </div>		<div>June 30th - Sochi</div> <div> <div>Netherland</div> <div>USA</div> </div>		<div>July 6th - Nizhny Novgorod</div> <div> <div>Netherland</div> <div>England</div> </div>		<div>July 10th - Saint Petersburg</div> <div> <div>Netherland</div> <div>France</div> </div>		<div>July 15th - Moscow</div> <div> <div>France</div> <div>Brazil</div> </div>	
<div> <div>A</div> <div>B</div> </div>		<div>June 30th - Kazan</div> <div> <div>Ecuador</div> <div>England</div> </div>		<div>July 6th - Kazan</div> <div> <div>Argentina</div> <div>France</div> </div>		<div>July 11th - Moscow</div> <div> <div>Belgium</div> <div>Brazil</div> </div>		<div>THIRD-PLACE GAME</div> <div> <div>Netherland</div> <div>Belgium</div> </div>	
<div> <div>C</div> <div>D</div> </div>		<div>July 2nd - Samara</div> <div> <div>Argentina</div> <div>Australia</div> </div>		<div>July 7th - Sochi</div> <div> <div>Spain</div> <div>Belgium</div> </div>					
<div> <div>C</div> <div>D</div> </div>		<div>July 2nd - Rostov-on-Don</div> <div> <div>Mexico</div> <div>France</div> </div>		<div>July 7th - Samara</div> <div> <div>Brazil</div> <div>Serbia</div> </div>					
<div> <div>E</div> <div>F</div> </div>		<div>July 1st - Moscow</div> <div> <div>Spain</div> <div>Croatia</div> </div>							
<div> <div>E</div> <div>F</div> </div>		<div>July 1st - Nizhny Novgorod</div> <div> <div>Germany</div> <div>Belgium</div> </div>							
<div> <div>G</div> <div>H</div> </div>		<div>July 3rd - Saint Petersburg</div> <div> <div>Brazil</div> <div>Korea</div> </div>							
<div> <div>G</div> <div>H</div> </div>		<div>July 3rd - Moscow</div> <div> <div>Serbia</div> <div>Portugal</div> </div>							

# 보완점

## 데이터 수집 및 전처리

- xG, xA, 시즌 별 패스 정확도 등 게임 데이터가 아닌 실제 데이터 활용  
→ 실제 경기 별 크롤링 / SKY SPORTS' 제공 선수 데이터 크롤링 / 실제 데이터를 어떻게 전처리해야 수치로 비교 가능할까?
- 경기 외적 요소 고려 필요  
→ 경기장 기온, 선수별 컨디션, 해발고도, 잔디 질, 관중 수, 공인구 종류 .. 등등

## 데이터 분석

- 각 팀별 취약지역과 우세지역 파악 → 전술 별 상성관계 파악
- 경기 점유율, 슈팅, 유효슈팅 등의 요소들이 경기에 어떤 영향을 주는지 분석 필요

# Q/A

FIFA WORLD CUP  
**QATAR**  
2022



The Official Ball of the World Cup QATAR 2022