

An analysis of the World Cup

Who were the best players?

Will Drake

A dark blue diagonal gradient bar that starts from the bottom left and extends towards the top right, covering the lower half of the slide.

Background Info:

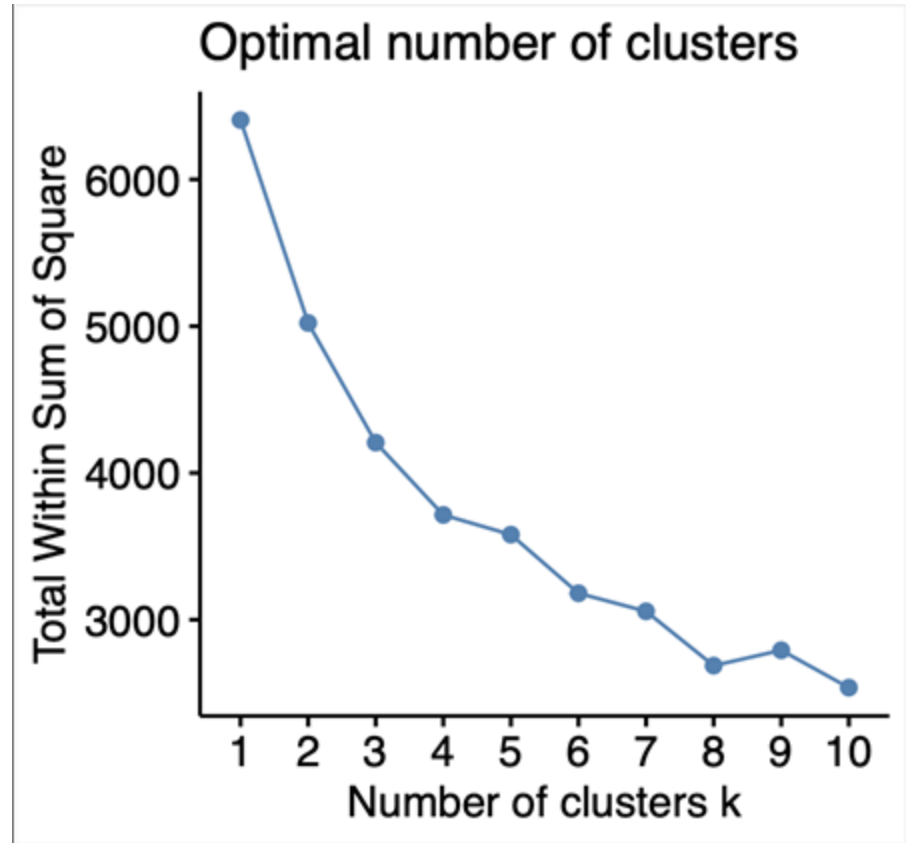
- Fifa World Cup held every 4 years
- 32 nations compete in a tournament
- In soccer there are 11 positions
 - Defenders
 - Goalkeeper
 - Fullbacks (Right and Left)
 - Center back
 - Sweeper
 - Midfielders
 - Defending
 - Center
 - Winger
 - Attacking
 - Forwards
 - Striker
 - Center Forward

Methods:

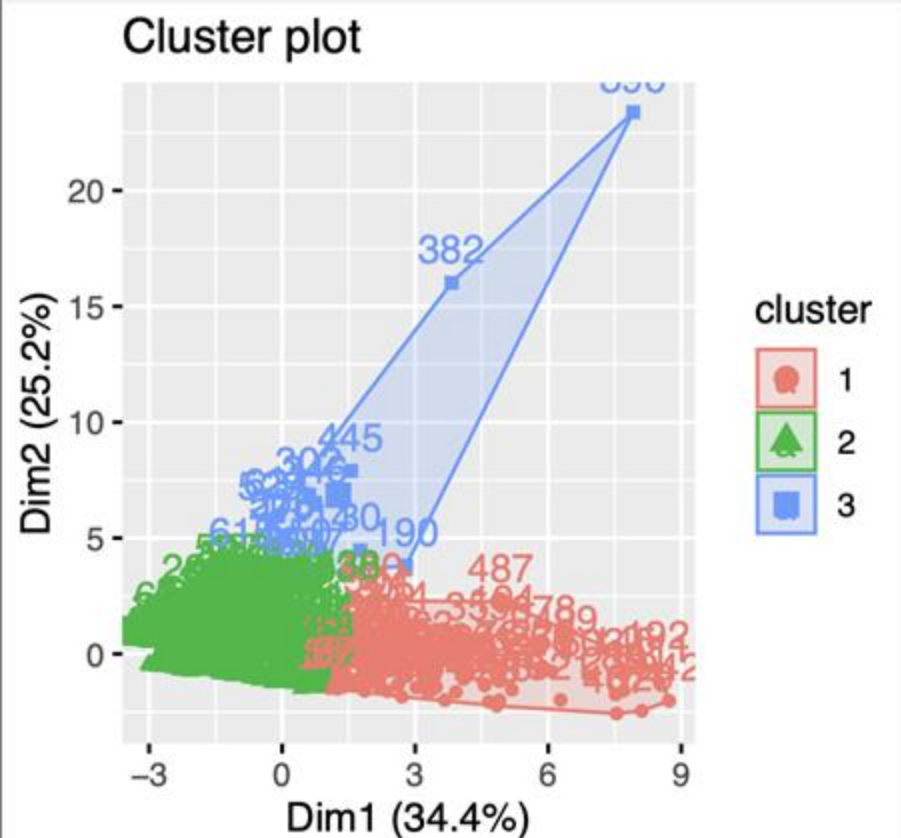
	Player	Position	Squad	Age	Born	Goals	Shots	Shots.on.Target	Average.Shot.Distance
1	Ali Abdi	DF	Tunisia	28	1993	0	0	0	NA
2	Salis Abdul Samad	MF	Ghana	22	2000	0	1	0	26.9
3	Saud Abdulhamid	DF	Saudi Arabia	22	1999	0	1	0	34.6
4	Nawaf Al Abed	MF	Saudi Arabia	32	1990	0	2	0	21.1
5	Pape Abou Cissé	DF	Senegal	26	1995	0	0	0	NA
6	Vincent Aboubakar	FW	Cameroon	30	1992	2	6	3	17.8
7	Abdulrahman Al-Aboud	MF	Saudi Arabia	27	1995	0	1	0	4.6
8	Zakaria Aboukhail	FW	Morocco	22	2000	1	3	2	14.9
9	Kellyn Acosta	MF	United States	26	1995	0	0	0	NA
10	Marcos Acuila	DF	Argentina	30	1991	0	2	0	17.9
11	Tyler Adams	MF	United States	23	1999	0	1	0	30.4
12	Sam Adegugbe	DF	Canada	27	1995	0	0	0	NA
13	Michel Aebischer	MF	Switzerland	25	1997	0	0	0	NA
14	Akram Alfi	FW	Qatar	25	1996	0	2	0	29.8
15	Nayef Aguerd	DF	Morocco	26	1996	0	1	0	8.0

- Data from <https://fbref.com/en/>
- Variables:
 - Player
 - Position
 - Squad
 - Age
 - Born
 - Goals
 - Shots, shots on target, & average shot distance
 - Free kicks made
 - Penalty kicks made & attempted
 - Touches
 - Tackles & tackles won
 - Blocks
 - Interceptions
 - Passes completed & attempted
- Added Rank Variable
- Created Subset
 - Removed NA's and some variables

K Means



Cluster



```
> prop.table(table(test_df$Position, test_df$cluster),margin = 1)
```

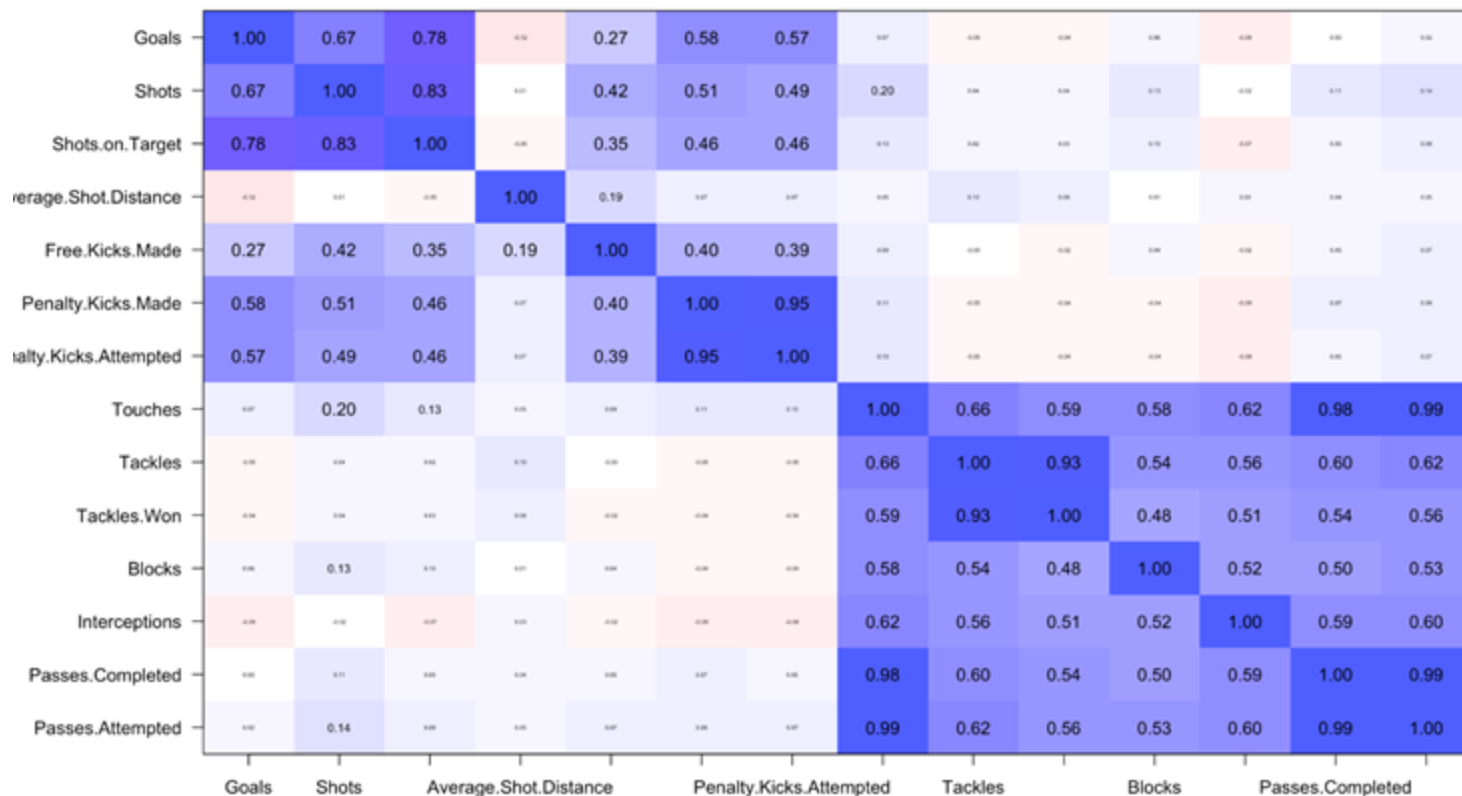
	1	2	3
DF	0.36923077	0.63076923	0.00000000
FW	0.04861111	0.86111111	0.09027778
MF	0.23376623	0.74675325	0.01948052

	model1			model2			model3			model4			model5			model6			model7		
Predictors	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p
(Intercept)	15.98	15.04 – 16.92	<0.001	17.33	16.12 – 18.55	<0.001	16.67	15.62 – 17.71	<0.001	12.52	10.25 – 14.79	<0.001	15.31	14.41 – 16.21	<0.001	15.20	14.32 – 16.08	<0.001	15.17	14.29 – 16.05	<0.001
Goals	-2.26	-3.27 – -1.25	<0.001																		
Shots				-0.67	-0.94 – -0.41	<0.001															
Shots on Target							-1.42	-1.97 – -0.87	<0.001												
Average Shot Distance										0.15	0.03 – 0.26	0.017									
Free Kicks Made													-2.00	-4.10 – 0.09	0.061						
Penalty Kicks Made																-2.98	-6.25 – 0.28	0.073			
Penalty Kicks Attempted																			-1.65	-4.26 – 0.96	0.214
Touches																					
Tackles																					
Tackles Won																					
Blocks																					
Interceptions																					
Passes Completed																					
Passes Attempted																					
Observations	428			428			428			428			428			428			428		
R ² / R ² adjusted	0.044 / 0.041			0.057 / 0.055			0.057 / 0.055			0.013 / 0.011			0.008 / 0.006			0.008 / 0.005			0.004 / 0.001		
AIC	3097.461			3091.485			3091.257			3110.804			3113.026			3113.327			3114.995		

Estimates	CI	p	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p		
19.33	18.05 – 20.60	<0.001	17.87	16.69 – 19.04	<0.001	17.47	16.37 – 18.57	<0.001	17.49	16.24 – 18.73	<0.001	17.08	15.96 – 18.21	<0.001	18.40	17.24 – 19.56	<0.001	18.76	17.54 – 19.97	<0.001		
																		-1.34	-2.99 – 0.30	0.109		
																		-0.18	-0.70 – 0.33	0.483		
																		-0.49	-1.58 – 0.61	0.383		
																		0.15	0.03 – 0.26	0.011		
																		-1.17	-3.38 – 1.03	0.296		
																		-5.09	-14.62 – 4.44	0.294		
																		6.84	-0.62 – 14.29	0.072		
-0.03	-0.04 – -0.02	<0.001																-0.05	-0.17 – 0.07	0.420		
			-0.73	-0.94 – -0.51	<0.001													0.03	-0.59 – 0.65	0.926		
						-1.07	-1.39 – -0.75	<0.001										-0.58	-1.38 – 0.23	0.158		
									-0.88	-1.22 – -0.55	<0.001							0.06	-0.40 – 0.52	0.799		
												-1.04	-1.43 – -0.65	<0.001				-0.14	-0.64 – 0.37	0.602		
															-0.03	-0.04 – -0.03	<0.001		-0.08	-0.17 – 0.01	0.083	
																	-0.03	-0.04 – -0.02	<0.001	0.10	-0.07 – 0.27	0.231
428			428			428			428			428			428			428				
0.144 / 0.142			0.093 / 0.091			0.092 / 0.090			0.059 / 0.057			0.062 / 0.060			0.128 / 0.126			0.132 / 0.129			0.229 / 0.203	
3050.204			3074.809			3075.389			3090.476			3089.283			3057.996			3056.200			3031.007	

Linear Regressions (multicollinearity)

Correlation plot from data




```
Call:
lm(formula = Goals ~ relevel(Position, ref = "FW"), data = Subset1)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.7986	-0.2273	-0.1538	0.2014	7.2014

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.79861	0.06656	11.998	< 2e-16 ***
relevel(Position, ref = "FW")DF	-0.64476	0.09663	-6.672	7.86e-11 ***
relevel(Position, ref = "FW")MF	-0.57134	0.09259	-6.171	1.59e-09 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7987 on 425 degrees of freedom

Multiple R-squared: 0.1153, Adjusted R-squared: 0.1112

F-statistic: 27.7 on 2 and 425 DF, p-value: 4.9e-12

```
Call:
lm(formula = Blocks ~ Position, data = Subset1)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.3615	-1.8750	-0.3615	1.1250	10.0195

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.3615	0.2151	15.627	< 2e-16 ***
PositionFW	-1.4865	0.2967	-5.010	8.01e-07 ***
PositionMF	-0.3810	0.2921	-1.304	0.193

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.453 on 425 degrees of freedom

Multiple R-squared: 0.06123, Adjusted R-squared: 0.05682

F-statistic: 13.86 on 2 and 425 DF, p-value: 1.474e-06

FW: Goals, Shots, Shot on target, Average Shot Distance MF: Passes completed, Passes attempted, and Touches
DF: Tackles, Tackles won, Blocks, Interceptions, Passes completed, Passes attempted, and Touches

Players	Posit... ¹	Goals	Shots	Shots... ²	Avera... ³	Free... ⁴	Penal... ⁵	Penal... ⁶	Touches	Tackles	Tackl... ⁷	Blocks	Inter... ⁸	Passe... ⁹	Passe... ^x	Team... ^x	BestFW
<chr>	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1 Kylian Mb...	FW	8.97	7.92	6.40	-0.0783	2.14	7.36	5.82	1.42	-0.937	-0.860	-1.08	-0.878	0.747	0.962	-1.43	23.4
2 Lionel Me...	FW	7.79	7.30	7.70	0.442	6.96	14.9	14.8	2.57	0.301	0.297	-0.287	-0.878	2.04	2.24	-1.54	22.4
3 Olivier G...	FW	4.25	3.91	3.17	-0.914	-0.276	-0.149	-0.161	-0.312	-0.737	-0.860	-0.683	-0.421	-0.524	-0.437	-1.43	12.2
4 Julián Ál...	FW	4.25	2.36	4.46	-0.750	-0.276	-0.149	-0.161	0.221	0.0412	-0.0833	2.09	-0.878	-0.0065	-0.0127	-1.54	11.8
5 Marc Rash...	FW	3.07	2.36	3.17	-0.147	4.55	-0.149	-0.161	-0.468	-0.737	-0.860	-0.287	-0.421	-0.514	-0.500	-0.000	8.75

1 Kylian Mbappé

2 Lionel Messi

3 Olivier Giroud

4 Julián Álvarez

5 Marc Rashford

Players	Posit... ¹	Goals	Shots	Shots... ²	Avera... ³	Free... ⁴	Penal... ⁵	Penal... ⁶	Touches	Tackles	Tackl... ⁷
<chr>	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1 Rodrigo Paul	MF	-0.469	1.13	1.22	0.332	-0.276	-0.149	-0.161	3.91	1.86	1.45
2 Luka Modrić	MF	-0.469	0.822	1.87	0.469	-0.276	-0.149	-0.161	3.81	3.16	3.00
3 Marcelo Bro...	MF	-0.469	0.205	0.574	-0.0378	-0.276	-0.149	-0.161	3.65	2.38	1.84
4 Aurélien Tc...	MF	0.712	0.822	0.574	0.825	-0.276	-0.149	-0.161	3.14	1.86	0.683
5 Enzo Fernán...	MF	0.712	1.44	1.22	0.990	-0.276	-0.149	-0.161	3.17	4.71	4.93

1 Rodrigo Paul

2 Luka Modrić

3 Marcelo Brozović

4 Aurélien Tchouaméni

5 Enzo Fernández

Players	Posit... ¹	Goals	Shots	Shots... ²	Avera... ³	Free... ⁴	Penal... ⁵	Penal... ⁶	Touche...	Tackles	Tackl... ⁷	Blocks	Inter... ⁸	Passe... ⁹	Passe... [×]	Team... [×]	BestDF
<chr>	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1 Achraf ...	DF	-0.469	0.205	-0.0742	1.21	2.14	-0.149	-0.161	2.73	5.75	5.70	1.69	3.24	2.16	2.34	-1.21	23.6
2 Nicolás...	DF	-0.469	-0.721	-0.722	-0.738	-0.276	-0.149	-0.161	4.04	1.60	1.45	1.30	3.24	4.40	4.10	-1.54	20.1
3 Joško G...	DF	0.712	-0.412	-0.0742	-0.128	-0.276	-0.149	-0.161	3.72	1.34	1.45	2.09	4.15	3.70	3.58	-1.32	20.0
4 Rodri	DF	-0.469	-0.721	-0.722	2.54	-0.276	-0.149	-0.161	4.61	0.301	-0.0882	0.109	0.493	5.51	5.12	-0.227	16.1
5 Theo He...	DF	0.712	0.513	0.574	-0.134	-0.276	-0.149	-0.161	1.94	3.16	4.15	0.901	0.951	1.93	1.96	-1.43	15.0

- 1 Achraf Hakimi
- 2 Nicolás Otamendi
- 3 Joško Gvardiol
- 4 Rodri
- 5 Theo Hernández

Next Steps

- 2018 World Cup
- Include Goalkeeper and go in depth with each position in defending, midfield, and forward
- Outliers
- Cluster based on Squad or other variables
- Include more variables to represent MF