FIFA 20 x Goals

FIFA20 Stats & Real Goals

Wonjae Lee

1/22/2021

Pitch

Predicting:

Number of goals in real life per minute based on FIFA attributes

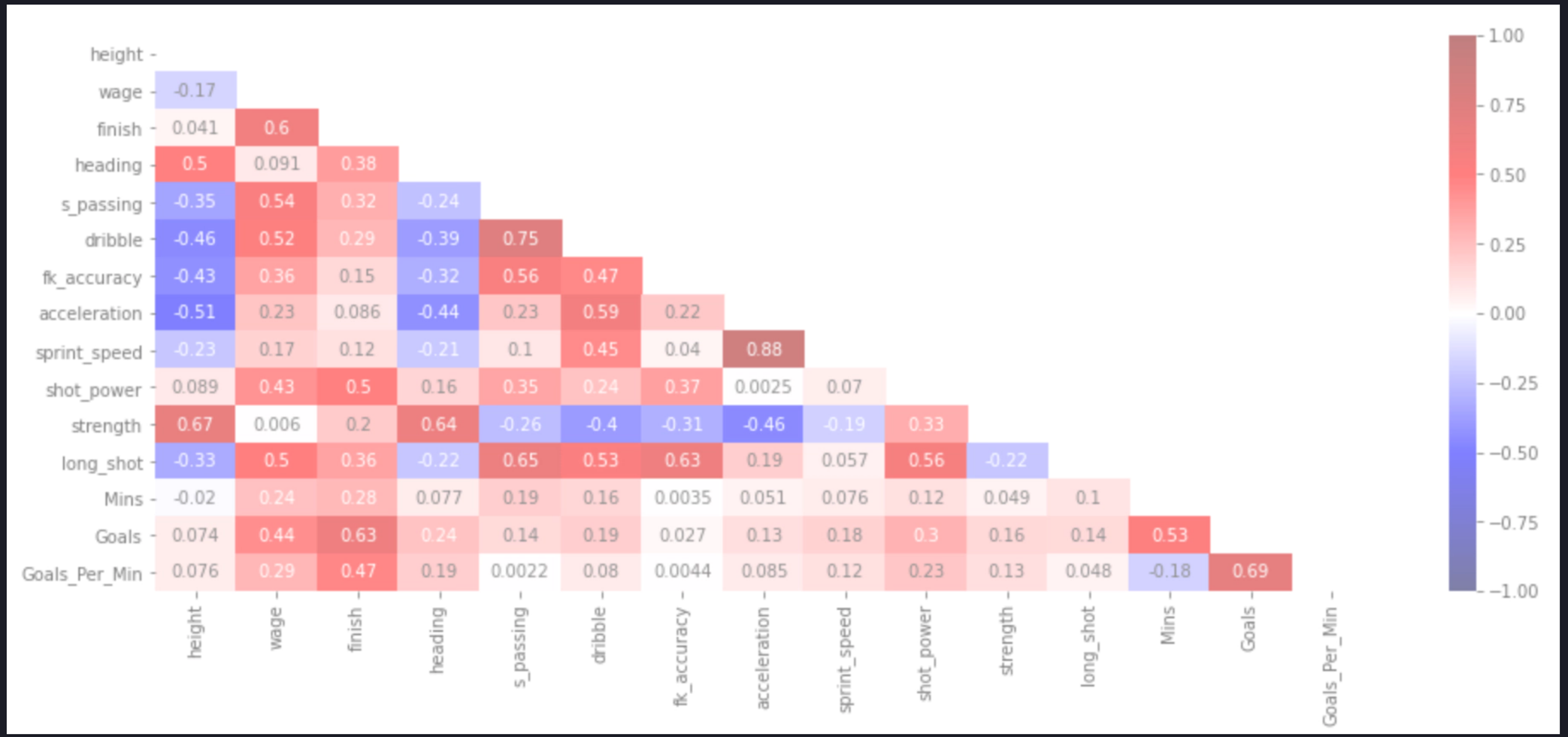
Pitch

Why is this important?

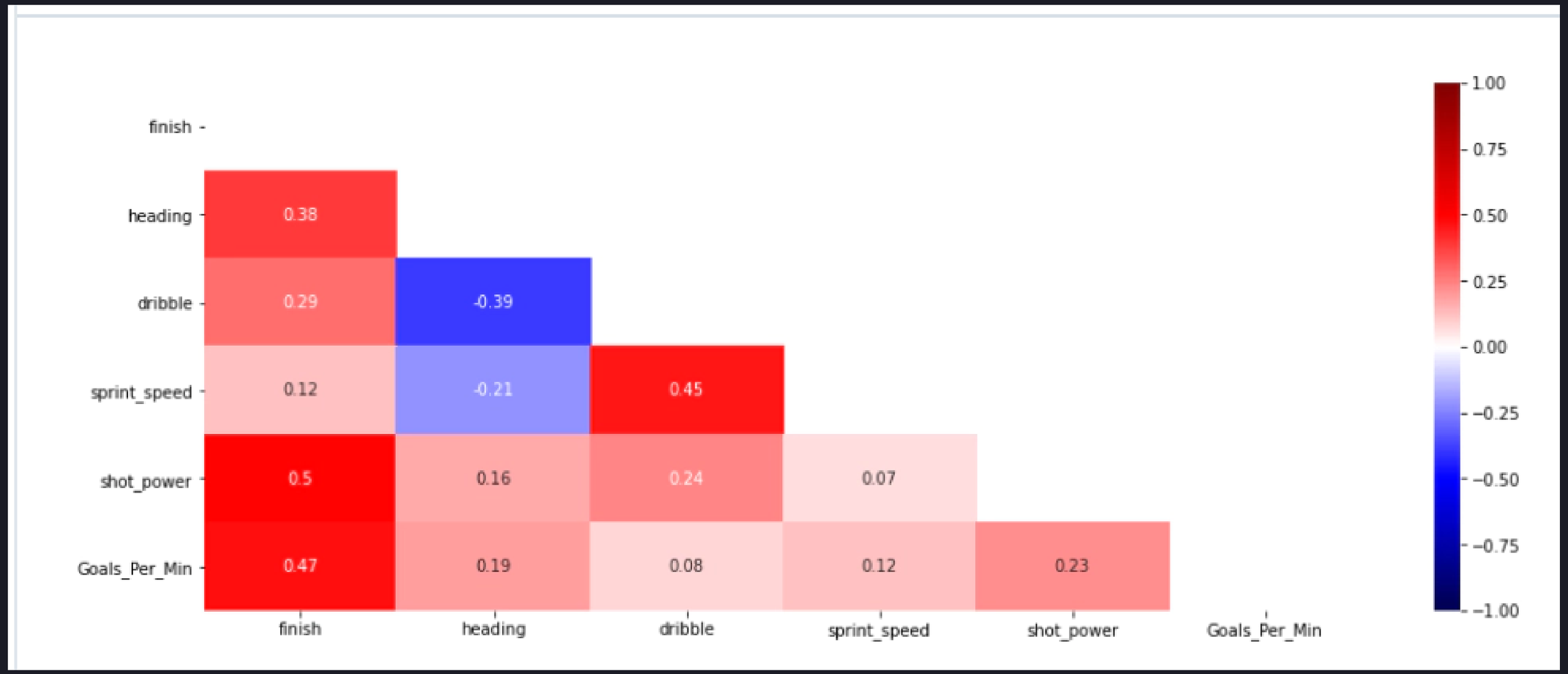
High-level details

- Can be a useful database for transfer target
- Can help teams understand the strength and weakness of players
- Validating the subjective nature of ratings

Pitch


height
wage -0.17
finish 0.041 0.6
heading 0.5 0.091 0.38
s_passing -0.35 0.54 0.32 -0.24
dribble -0.46 0.52 0.29 -0.39 0.75
fk_accuracy -0.43 0.36 0.15 -0.32 0.56 0.47
acceleration -0.51 0.23 0.086 -0.44 0.23 0.59 0.22
sprint_speed -0.23 0.17 0.12 -0.21 0.1 0.45 0.04 0.88
shot_power 0.089 0.43 0.5 0.16 0.35 0.24 0.37 0.0025 0.07
strength 0.67 0.006 0.2 0.64 -0.26 -0.4 -0.31 -0.46 -0.19 0.33
long_shot -0.33 0.5 0.36 -0.22 0.65 0.53 0.63 0.19 0.057 0.56 -0.22
Mins -0.02 0.24 0.28 0.077 0.19 0.16 0.0035 0.051 0.076 0.12 0.049 0.1
Goals 0.074 0.44 0.63 0.24 0.14 0.19 0.027 0.13 0.18 0.3 0.16 0.14 0.53
Goals_Per_Min 0.076 0.29 0.47 0.19 0.0022 0.08 0.0044 0.085 0.12 0.23 0.13 0.048 -0.18 0.69
height
wage
finish
heading
s_passing
dribble
fk_accuracy
acceleration
sprint_speed
shot_power
strength
long_shot
Mins
Goals
Goals_Per_Min
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Pitch


finish
heading
dribble
sprint_speed
shot_power
Goals_Per_Min
0.38
0.29
-0.39
0.12
-0.21
0.45
0.5
0.16
0.24
0.07
0.47
0.19
0.08
0.12
0.23
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Pitch

Ordinary Least Squared

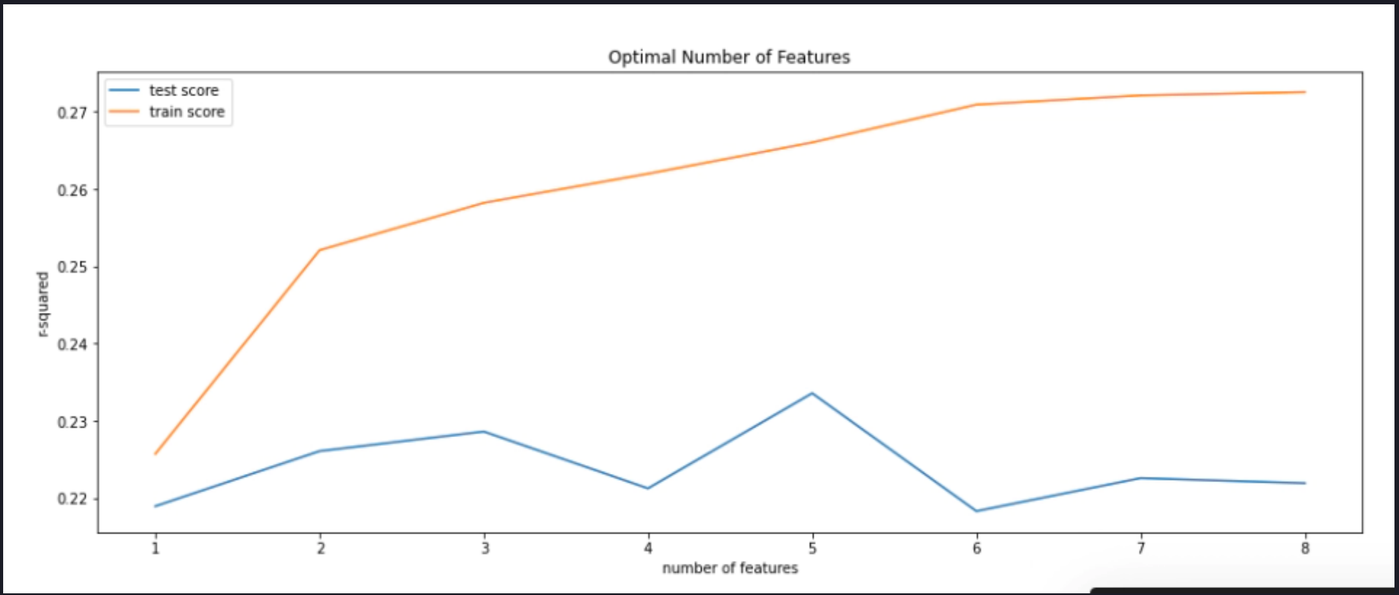
OLS Regression Results

Dep. Variable:	Goals_Per_Min	**R-squared:**	0.240
Model:	OLS	**Adj. R-squared:**	0.225
Method:	Least Squares	**F-statistic:**	16.37
Date:	Fri, 22 Jan 2021	**Prob (F-statistic):**	4.91e-14
Time:	07:26:09	**Log-Likelihood:**	1305.9
No. Observations:	265	**AIC:**	-2600.
Df Residuals:	259	**BIC:**	-2578.
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|
const	-0.0104	0.002	-4.760	0.000
finish	0.0002	3.03e-05	6.932	0.000
heading	-5.469e-06	1.18e-05	-0.464	0.643
dribble	-3.867e-05	2.17e-05	-1.785	0.075
sprint_speed	2.331e-05	1.23e-05	1.897	0.059
shot_power	5.285e-07	2.51e-05	0.021	0.983

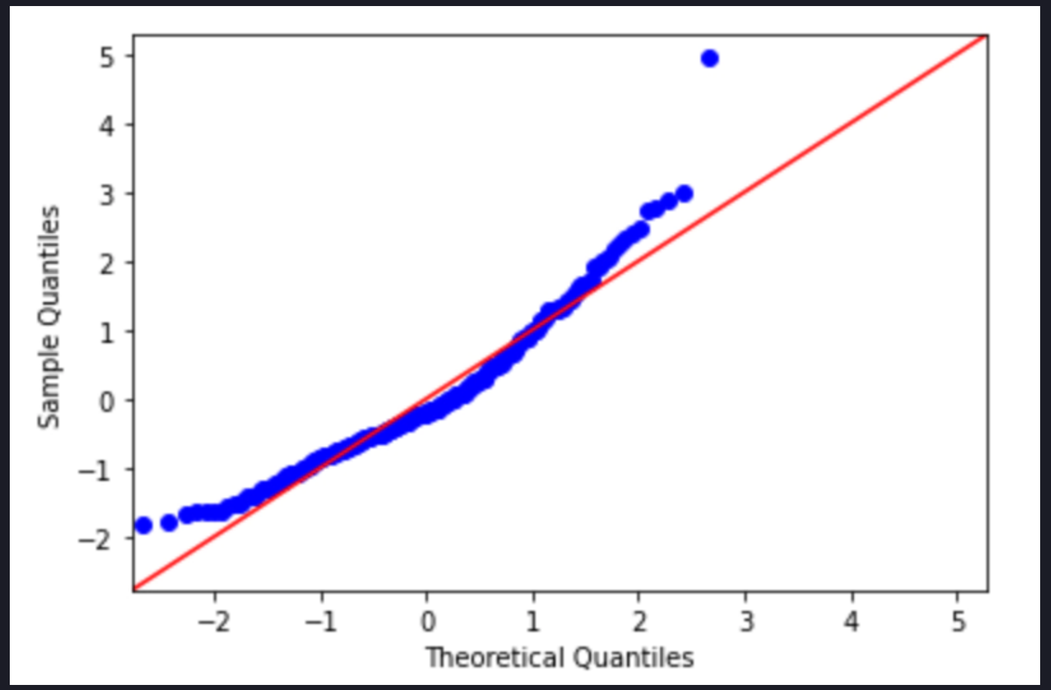
Pitch

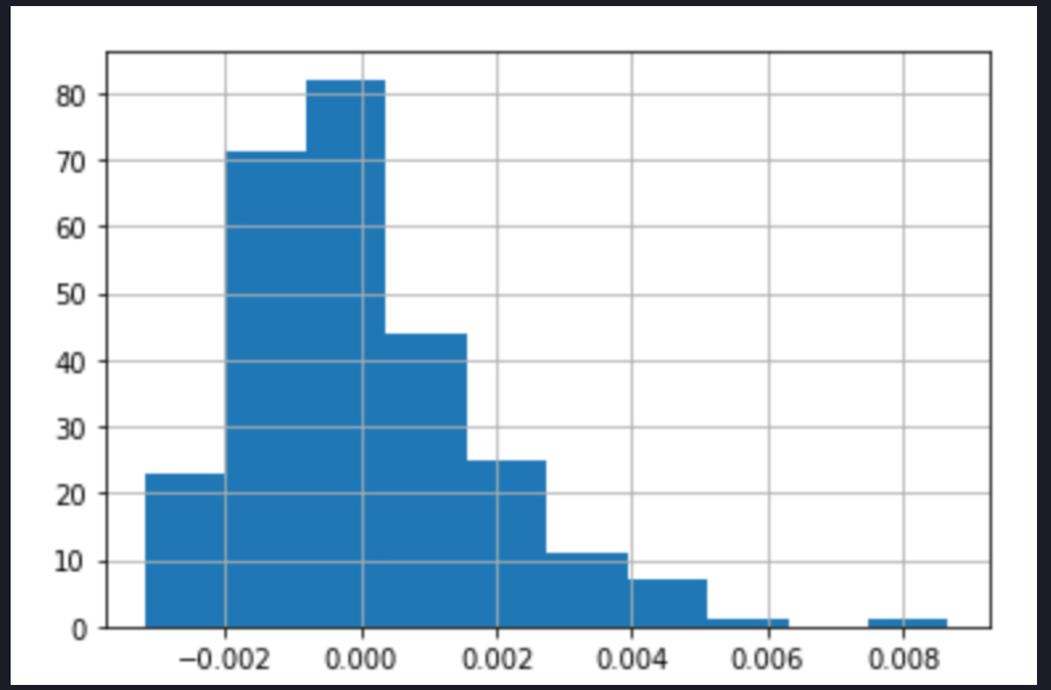
Grid Search Cross Validation



param_n_features	rank_test_score	mean_test_score	mean_train_score
1	7	0.218945	0.225720
2	3	0.226075	0.252079
3	2	0.228610	0.258209
4	6	0.221253	0.261980
5	1	0.233566	0.266023
6	8	0.218308	0.270929
7	4	0.222586	0.272124
8	5	0.221927	0.272547

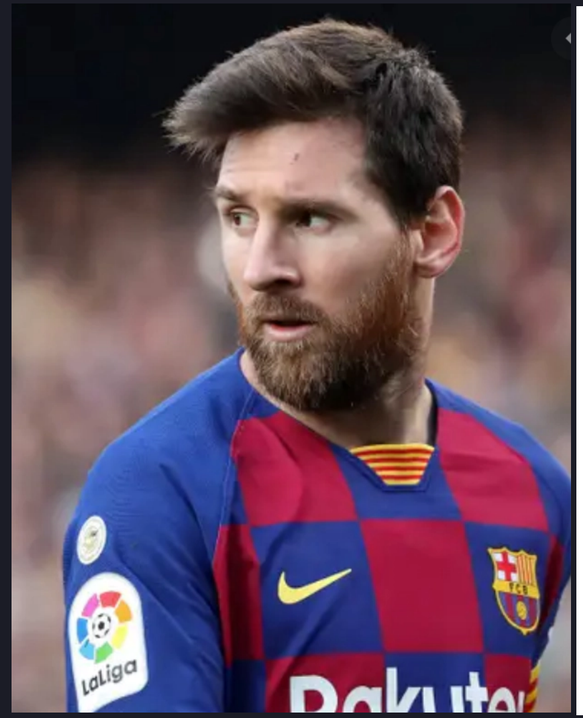
Pitch


Sample Quantiles
Theoretical Quantiles
5
4
3
2
1
0
−1
−2
−2
−1
0
1
2
3
4
5


80
70
60
50
40
30
20
10
0
−0.002
0.000
0.002
0.004
0.006
0.008

Pitch

height	70
pf	Right
wage	69
finish	78
heading	65
s_passing	75
dribble	85
fk_accuracy	60
acceleration	86
sprint_speed	90
shot_power	80
strength	75
long_shot	66
work_rate	High
Mins	1370
Goals	18
Goals_Per_Min	0.0131387



height	67
pf	Left
wage	560
finish	95
heading	70
s_passing	92
dribble	97
fk_accuracy	94
acceleration	91
sprint_speed	84
shot_power	86
strength	68
long_shot	94
work_rate	Medium
Mins	3067
Goals	25
Goals_Per_Min	0.00815129

```
y_pred = baseline_fit.predict(X.loc[[' L. Muriel']])
print('Actual goal per min:', y.loc[' L. Muriel'])
print('Predicted goal per min:', y_pred[0])
print('Residual:', y.loc[' L. Muriel']-y_pred[0])
```

```
Actual goal per min: 0.013138686
Predicted goal per min: 0.004475661085706088
Residual: 0.008663024914293913
```

Pitch

What we learned

There were several limitations and short-comings in our model in predicting the number of goals per minute based on FIFA attributes

---

- Many features were collinear (strength-height, speed-acceleration, etc.
- Try different regressions like Lasso, Ridge, etc.
- Adjust by level of competitiveness of the leagues
- Look at the data over multiple seasons

Thank you

Pitch