



English Premier League 2018-2019

Nikolas Rakryan Widagdo

Introduction

This data set is wide ranging in the sense it encompass stats seen on a regular league table but goes beyond looking at how teams pass and keep possession, how they defend, tackle as well as looking at market values of a team and how much money each team was allotted from the TV rights deal.

This data was gathered from

- 1) BBC Sports Football,
- 2) Premierleague.com
- 3) Transfermarkt.co.uk

This data was not scrapped in a conventional sense and appears in a rather haphazard manner. To counter this I included category descriptors at the start of each variable name, this should help to provide a more cohesive understanding of the data set as well as aid in sub setting.

Data pre-processing

Pre-processing done:

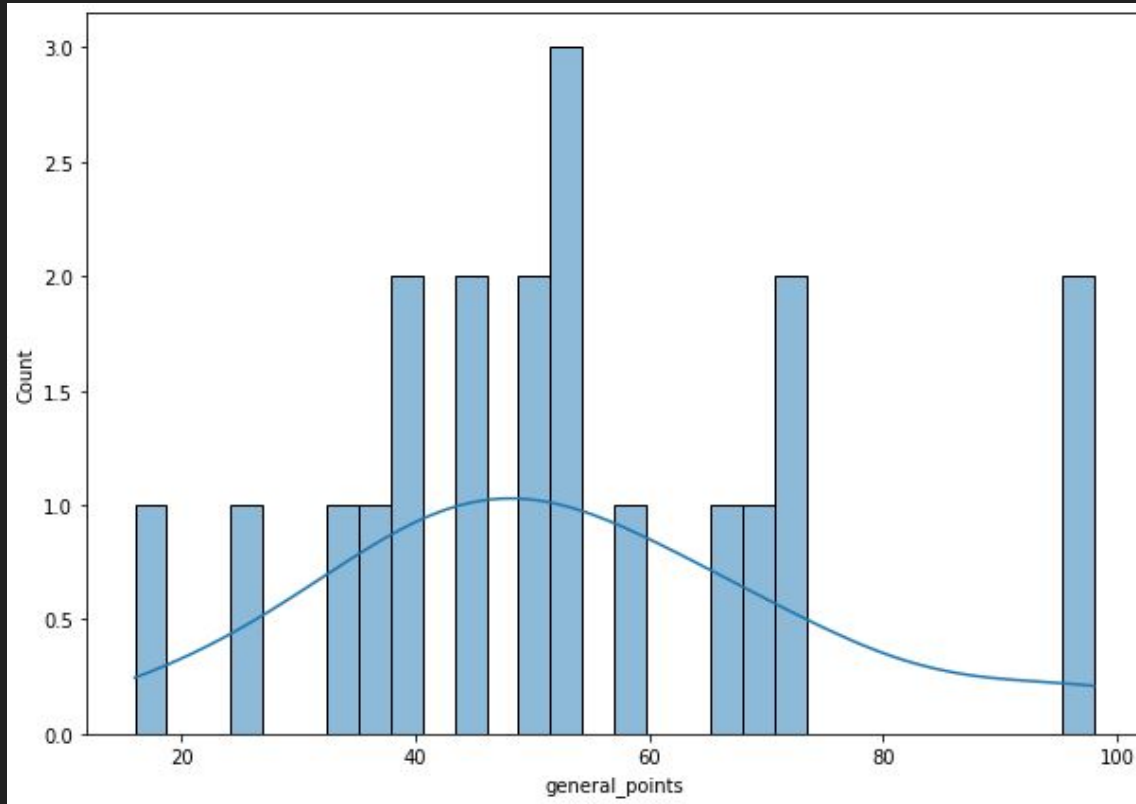
- Checked for duplicate data
- Erased white spaces between column names
- Erased white spaces for Bournemouth in Team column
- Checked for missing values
- Checking the data type
- Assigning new data type for incorrect data types

How clean is the data ?

The dataset is fairly clean:

- No missing values
- No duplicates
- No outliers

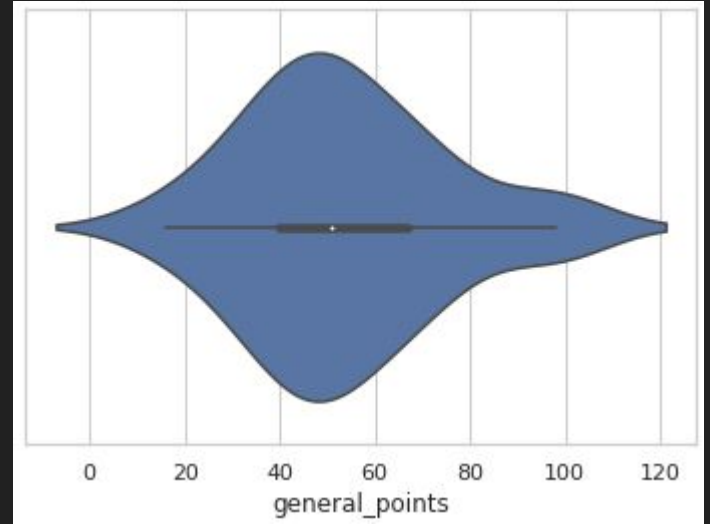
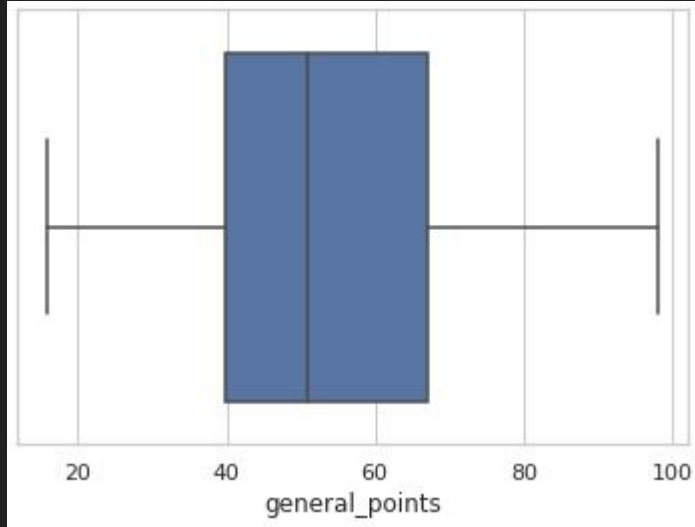
Point distribution of EPL teams?



The point distribution of EPL is reasonably distributed normally.

No apparent outliers in the data.

Point distribution of EPL teams?



Upon investigation with other plots, it is evident that there were no outliers.

Best attacking team ?

Team	attack_mean_value
Manchester City	2332.200000
Chelsea	2181.133333
Liverpool	2092.266667
Tottenham	1904.133333
Arsenal	1843.066667

The top 5 and bottom 5 teams based on their attack_mean_value can be seen here.

Based on the analysis it is proven that the best attacking team is Manchester City.

Meanwhile for the worst attacking team goes to Cardiff.

Team	attack_mean_value
Cardiff	1036.266667
Burnley	1293.666667
Newcastle	1319.866667
Brighton	1345.266667
Southampton	1358.000000

Best defending team ?

Team	defence_mean_value
Newcastle	393.000
Southampton	386.375
Burnley	367.375
Cardiff	363.625
Brighton	357.375

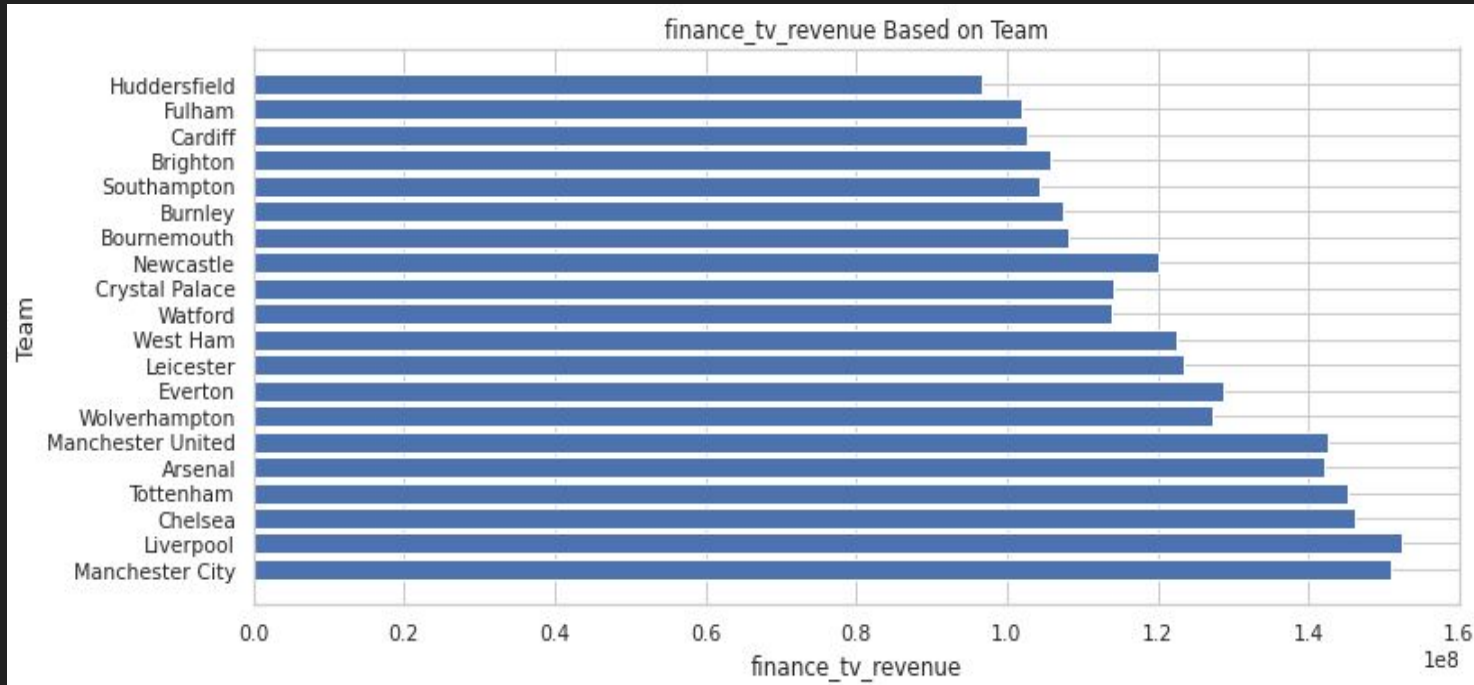
Team	defence_mean_value
Manchester City	230.375
Liverpool	254.625
Chelsea	256.750
Manchester United	293.625
Tottenham	296.750

The top 5 and bottom 5 teams based on their defence_mean_value can be seen here.

Based on the analysis it is proven that the best defending team is Newcastle.

Meanwhile for the worst attacking team goes to Manchester City.

Team with good financial aspect?



Here we can see the finance_tv_revenue based on teams.

Team with good financial aspect?

Team	finance_tv_revenue
Liverpool	152425146
Manchester City	150986355
Chelsea	146030216
Tottenham	145230801
Manchester United	142512868

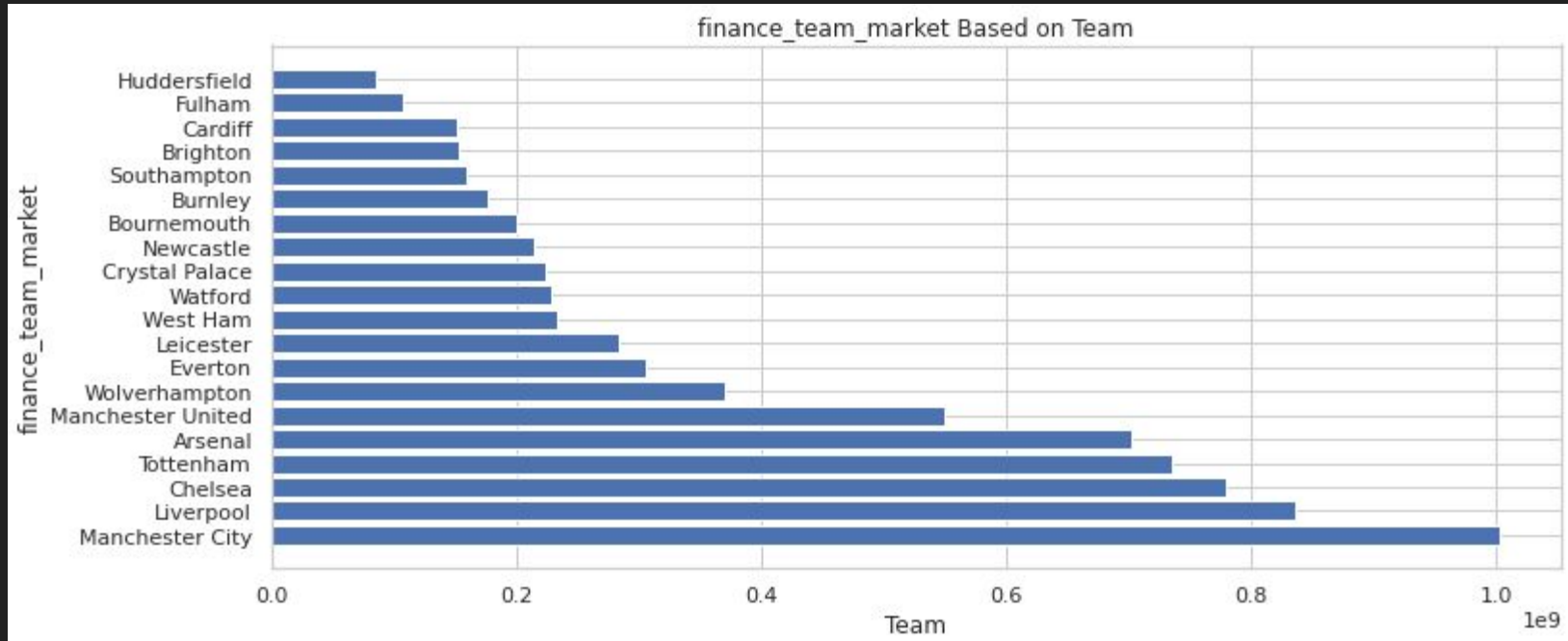
The top 5 and bottom 5 teams based on their finance_tv_revenue can be seen here.

Based on the analysis it is proven that the team with the best finance_tv_revenue is Liverpool.

Meanwhile for the team with the worst finance_tv_revenue goes to Huddersfield.

Team	finance_tv_revenue
Huddersfield	96628865
Fulham	101904692
Cardiff	102704107
Southampton	104302937
Brighton	105741728

Team with good financial aspect?



Here we can see the finance_team_market based on teams.

Team with good financial aspect?

Team	finance_team_market
Manchester City	1003200000
Liverpool	836440000
Chelsea	779460000
Tottenham	735240000
Arsenal	701800000

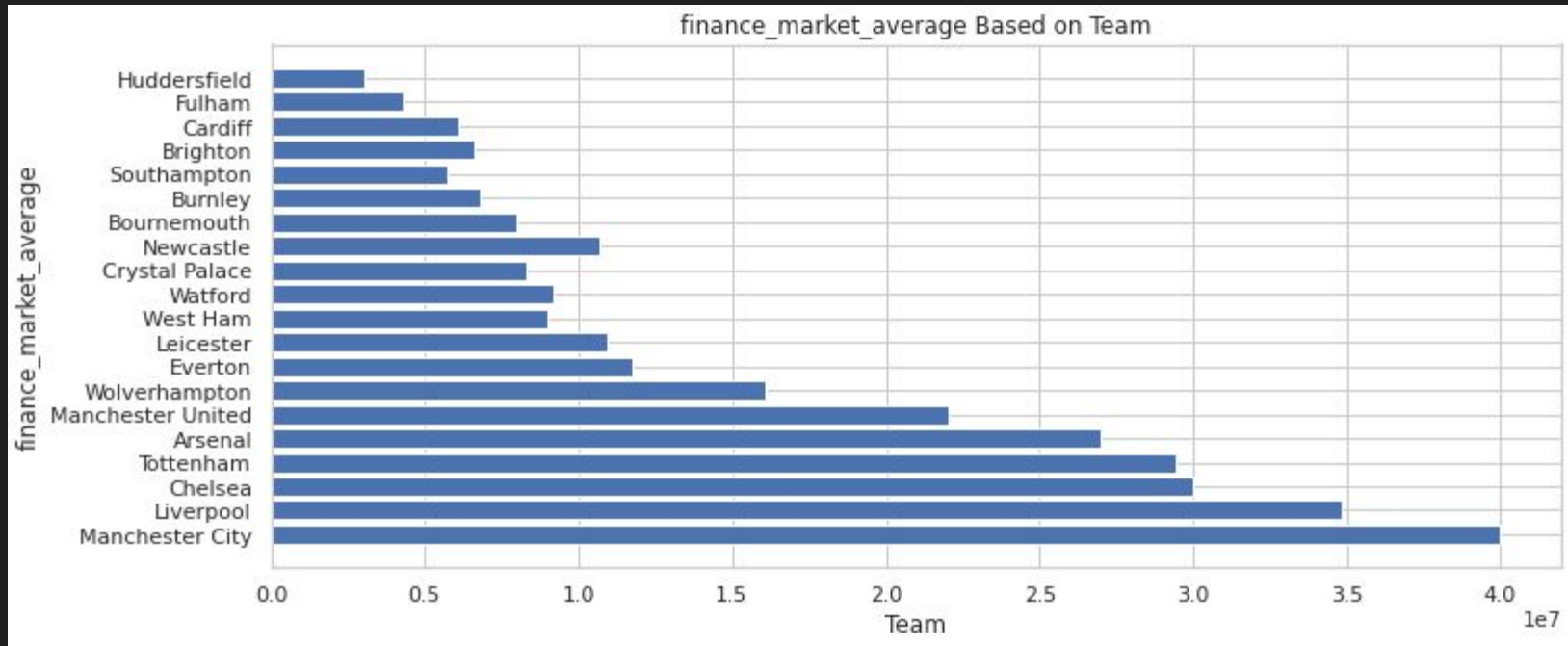
The top 5 and bottom 5 teams based on their finance_team_market can be seen here.

Based on the analysis it is proven that the team with the best finance_team_market is Manchester City.

Meanwhile for the team with the worst finance_team_market goes to Huddersfield.

Team	finance_team_market
Huddersfield	85492000
Fulham	106920000
Cardiff	152020000
Brighton	152680000
Southampton	160072000

Team with good financial aspect?



Here we can see the finance_market_average based on teams.

Team with good financial aspect?

Team	finance_market_average
Manchester City	39987200
Liverpool	34848000
Chelsea	29981600
Tottenham	29409600
Arsenal	26989600

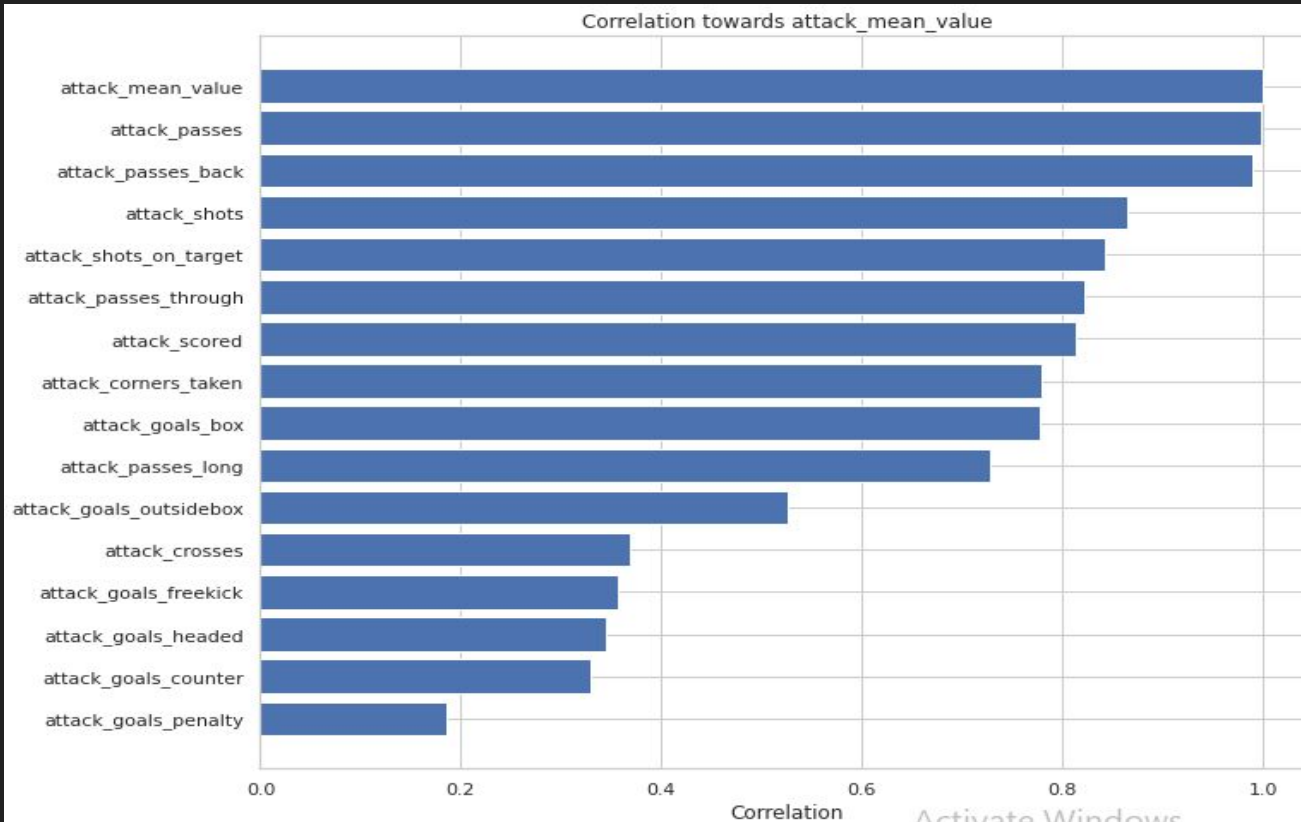
Team	finance_market_average
Huddersfield	3053600
Fulham	4276800
Southampton	5720000
Cardiff	6080800
Brighton	6635200

The top 5 and bottom 5 teams based on their finance_market_average can be seen here.

Based on the analysis it is proven that the team with the best finance_market_average is Manchester City.

Meanwhile for the team with the worst finance_market_average goes to Huddersfield.

Other insights?



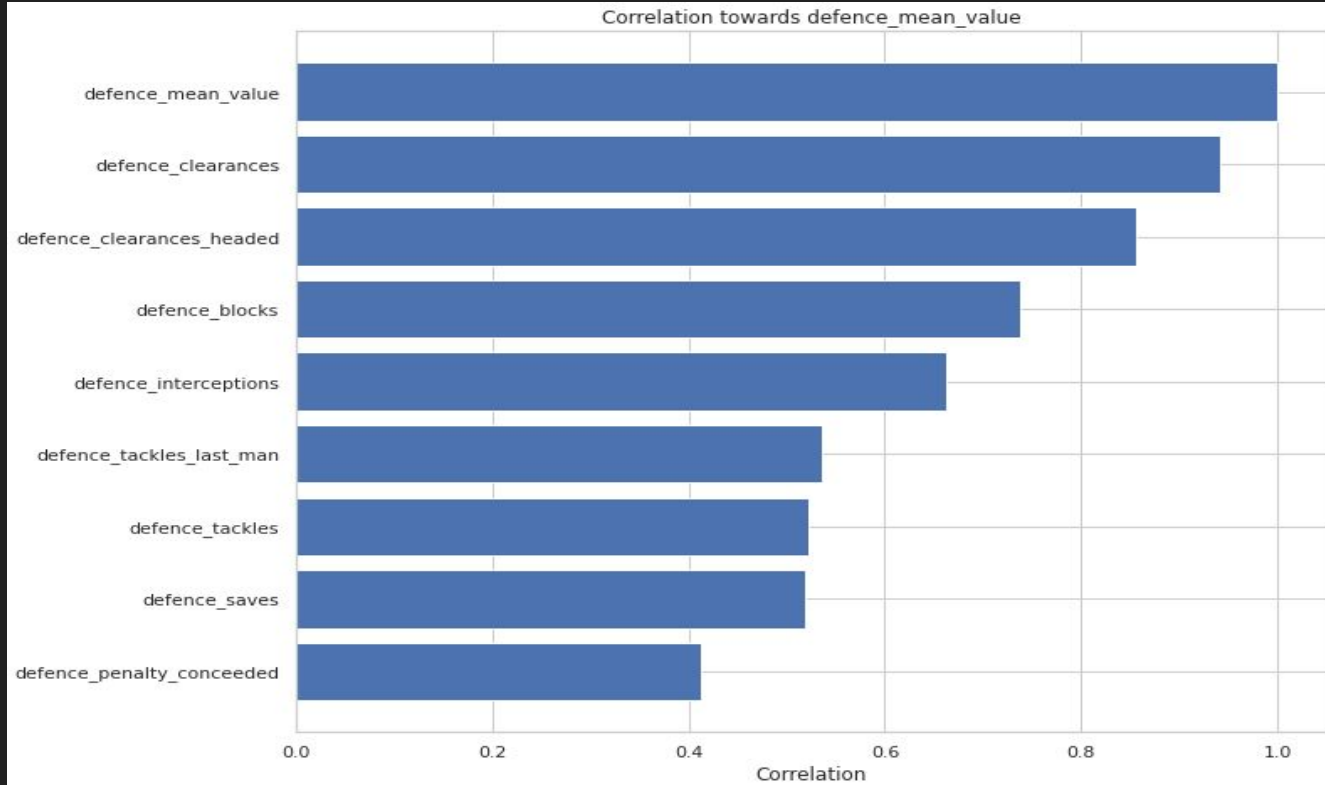
Here is presented the correlation level of features that make up the attack_mean_value.

Other insights?

attack_mean_value	1.000000
attack_passes	0.998567
attack_passes_back	0.990340
attack_shots	0.864091
attack_shots_on_target	0.841946
attack_passes_through	0.821999
attack_scored	0.813074
attack_corners_taken	0.779648
attack_goals_box	0.778425
attack_passes_long	0.727419
attack_goals_outsidebox	0.526448
attack_crosses	0.369397
attack_goals_freekick	0.356720
attack_goals_headed	0.344903
attack_goals_counter	0.329918
attack_goals_penalty	0.185614

The feature `attack_passes` has the highest correlation with `attack_mean_value`.

Other insights?



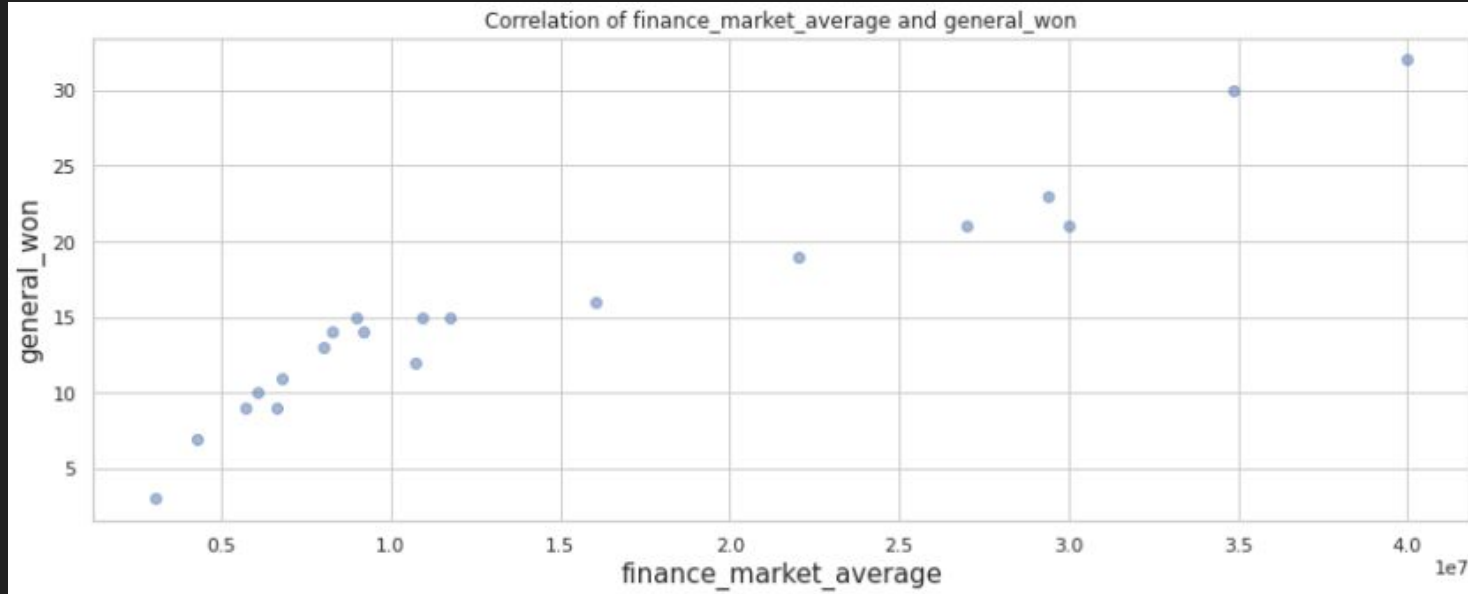
Here is presented the correlation level of features that make up the defend_mean_value.

Other insights?

defence_mean_value	1.000000
defence_clearances	0.941694
defence_clearances_headed	0.855709
defence_blocks	0.737325
defence_interceptions	0.662586
defence_tackles_last_man	0.535127
defence_tackles	0.522327
defence_saves	0.518176
defence_penalty_conceded	0.412019

The feature defence_clearance has the highest correlation with defence_mean_value.

Other insights?



This is a correlation plot of finance_market_average with general_won. We can say that the bigger the finance_market_average is, the bigger general_won can get.

Thank you

