

newest

March 23, 2022

```
[435]: import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.decomposition import PCA
from sklearn.tree import DecisionTreeClassifier
from sklearn.linear_model import LogisticRegression, Lasso, LassoCV
import os
from functools import reduce
from sklearn.metrics import r2_score
from sklearn.model_selection import cross_val_score
from sklearn.preprocessing import StandardScaler
from scipy.stats import norm
```

```
[260]: dfs = {}
for file in os.listdir("player_data"):
    dfs[file[:-4]] = pd.read_csv("player_data/"+file)
```

```
[384]: tournament_dfs_2021 = []
for df in ["tournament_shooting", "tournament_passing", "tournament_defense",
    ↪ "tournament_goalkeeping"]:
    tournament_dfs_2021.append(dfs[df][dfs[df].Year==2021])

merged_df = reduce(lambda left, right: pd.merge(left, right,
    ↪                                     how='outer',
    ↪                                     on=["Player", "Nation", "Age", "Pos", "Born", "League", "Year"]),
    tournament_dfs_2021)
```

```
[262]: leagues_dfs_2021 = []
for df in ["league_shooting", "league_passing", "league_defense",
    ↪ "league_goalkeeping"]:
    leagues_dfs_2021.append(dfs[df][dfs[df].Year==2021])

leagues_merged_df = reduce(lambda left, right: pd.merge(left, right,
    ↪                                     how='outer',
```

```

        ↪ on=["Player", "Nation", "Age", "Pos", "Born", "League", "Year"]),
        leagues_dfs_2021)

```

```

# leagues_merged_df = leagues_merged_df.drop("Nation", axis=1)

```

```

[263]: # merged_df = merged_df[["Player", "Nation"]]

```

```

[34]: # league_tourn_merged = pd.merge(left=merged_df, right=leagues_merged_df,
        ↪ on="Player", how="left")

```

```

[35]: # Use League data
# merged_df = league_tourn_merged

```

```

[264]: df = merged_df.groupby("Nation").mean()
df["Number Players"] = merged_df.groupby("Nation")["Player"].count()
df = df.dropna(how="all", axis=1)
df["TOP10"] = df.index.isin(dfs["tournament_results"]["2021"][:10].dropna()).
        ↪ astype("int")
df = df.drop("Dosqaly") # No GK
df = df.drop(["Year", "Penalty Kicks Save%"], axis=1)

```

```

[265]: # useful_cols = ["Age", "Gls", "Standard Sh/90", "Standard SoT/90", "Standard G/
        ↪ SoT",
#
        "Performance PK", "Performance GA90", "Performance Save%",
        ↪ "Pressures %",
#
        "W", "Performance CS%", "Total Cmp%",
#
        "Penalty Kicks PKA", "TOP10"]
# df = df[useful_cols]

```

```

[266]: # sns.pairplot(df, hue="TOP10")

```

```

[267]: mapping = {}
for i, country in enumerate(dfs["tournament_results"]["2021"].dropna()):
    mapping[country] = i+1
df.loc[:, "Score"] = df.index.map(mapping)
df.loc[:, "Score"] = df.loc[:, "Score"].fillna(0)

```

```

[268]: df

```

```

[268]:

```

	Age	Born	90s_x	Gls \
Nation				
Bernepamar	25.952381	1994.047619	2.776667	0.298571
Byasier Pujan	26.727273	1993.409091	2.086667	0.104286
Djipines	24.944444	1995.166667	2.435000	0.043889
Eastern Niasland	26.857143	1993.190476	1.594762	0.021905
Eastern Sleboube	27.100000	1993.000000	1.637500	0.051000

Esia	26.789474	1993.421053	2.301579	0.103158
Galamily	26.714286	1993.428571	2.587143	0.091905
Giumle Lizeibon	27.136364	1992.909091	2.175455	0.141364
Greri Landmoslands	27.368421	1992.631579	2.244211	0.085263
Ledian	27.800000	1992.300000	1.644000	-0.000500
Leoneku Guidisia	27.421053	1992.578947	1.748947	0.012632
Manlisgamncent	26.857143	1993.285714	2.263810	0.220000
Mico	26.285714	1993.761905	3.291429	0.182857
New Uwi	27.666667	1992.555556	1.831111	0.037778
Nganion	25.523810	1994.714286	3.684286	0.127619
Ngoque Blicri	25.700000	1994.300000	1.631500	0.018500
Nkasland Cronestan	26.052632	1994.105263	1.715789	0.116316
People's Land of Maneau	24.333333	1995.761905	4.027619	0.090000
Quewenia	27.000000	1993.050000	3.097000	0.111500
Sobianitedrucy	26.840000	1993.240000	3.510400	0.135600
Southern Ristan	28.125000	1991.958333	2.299167	0.089167
Varijitri Isles	24.444444	1995.611111	1.849444	0.024444
Xikong	28.315789	1991.789474	2.463684	0.105263

	Standard Sh	Standard SoT	Standard SoT%	\
Nation				
Bernepamar	1.092857	0.516190	36.777500	
Byasier Pujan	1.053810	0.251905	28.472143	
Djipines	0.690000	0.180556	24.854167	
Eastern Niasland	1.351905	0.166667	23.220000	
Eastern Sleboube	0.580000	0.098500	27.222222	
Esia	1.135789	0.318947	35.269286	
Galamily	1.230000	0.347619	22.672500	
Giumle Lizeibon	1.825909	0.963182	30.498000	
Greri Landmoslands	1.592105	0.433684	43.400833	
Ledian	2.155500	0.027500	10.913636	
Leoneku Guidisia	0.744211	0.068947	17.377500	
Manlisgamncent	1.197619	0.386667	31.392000	
Mico	1.614762	0.589048	38.660000	
New Uwi	0.428889	0.203889	49.580000	
Nganion	1.818095	0.632857	32.285000	
Ngoque Blicri	1.994000	0.303500	24.137333	
Nkasland Cronestan	1.714211	0.304737	21.028462	
People's Land of Maneau	0.627143	0.222857	37.751538	
Quewenia	1.217500	0.402500	24.178125	
Sobianitedrucy	1.518800	0.400400	29.788947	
Southern Ristan	0.726250	0.261667	29.703125	
Varijitri Isles	1.265556	0.371667	26.496429	
Xikong	0.786316	0.211579	28.376154	

	Standard Sh/90	Standard SoT/90	Standard G/Sh	...	\
Nation				...	

Bernepamar	1.226667	0.571905	1.441667e-01	...
Byasier Pujan	1.034286	0.244762	6.428571e-02	...
Djipines	0.704444	0.170556	8.583333e-02	...
Eastern Niasland	1.266667	0.146190	1.000000e-01	...
Eastern Sleboubé	0.583000	0.099500	3.111111e-02	...
Esia	1.111579	0.310526	4.571429e-02	...
Galamily	1.268095	0.340952	2.937500e-02	...
Giumle Lizeibon	2.025909	1.146364	6.666667e-02	...
Greri Landmoslands	1.591579	0.485789	7.083333e-02	...
Ledian	2.563500	0.014500	6.545455e-02	...
Leoneku Guidisia	0.713684	0.053684	-3.469447e-18	...
Manlisgamncent	1.214762	0.374762	1.246667e-01	...
Mico	1.660476	0.630952	9.812500e-02	...
New Uwi	0.418889	0.201667	2.133333e-01	...
Nganion	1.826190	0.641429	5.888889e-02	...
Ngoque Blicri	1.897500	0.308000	4.200000e-02	...
Nkasland Cronestan	1.524211	0.327368	5.076923e-02	...
People's Land of Maneau	0.590476	0.226667	1.715385e-01	...
Quewenia	1.223500	0.388500	1.068750e-01	...
Sobianitedrucy	1.542800	0.418400	1.010526e-01	...
Southern Ristan	0.712917	0.275000	6.500000e-02	...
Varijitri Isles	1.260556	0.356667	4.571429e-02	...
Xikong	0.783158	0.212105	3.692308e-02	...

	L	Performance	CS	Performance	CS%	\
Nation						
Bernepamar	0.530		0.080		0.020	
Byasier Pujan	0.270		0.580		50.000	
Djipines	0.550		0.300		25.070	
Eastern Niasland	0.990		-0.050		-0.080	
Eastern Sleboubé	0.805		0.230		25.045	
Esia	0.540		0.040		-0.040	
Galamily	0.500		0.300		40.040	
Giumle Lizeibon	0.550		0.270		24.930	
Greri Landmoslands	0.480		0.290		24.930	
Ledian	0.630		-0.020		-0.040	
Leoneku Guidisia	0.640		0.270		33.330	
Manlisgamncent	0.540		0.010		0.100	
Mico	0.540		0.070		16.740	
New Uwi	0.270		-0.010		-0.100	
Nganion	-0.050		0.290		33.380	
Ngoque Blicri	0.710		0.410		33.310	
Nkasland Cronestan	0.670		-0.040		-0.040	
People's Land of Maneau	-0.070		0.580		71.400	
Quewenia	0.260		-0.060		0.000	
Sobianitedrucy	-0.020		0.645		42.960	
Southern Ristan	0.220		0.620		59.920	

Varijitri Isles	0.960	-0.090	-0.020
Xikong	0.320	0.440	50.090

	Performance PKatt_y	Penalty Kicks PKA	\
Nation			
Bernepamar	0.13	0.08	
Byasier Pujan	-0.04	0.09	
Djipines	0.07	0.07	
Eastern Niasland	0.41	-0.10	
Eastern Sleboubé	-0.09	0.00	
Esia	-0.09	-0.04	
Galamily	-0.04	0.04	
Giumle Lizeibon	0.31	0.16	
Greri Landmoslands	0.21	0.22	
Ledian	0.68	0.39	
Leoneku Guidisia	0.28	-0.05	
Manlisgamncent	0.23	0.27	
Mico	0.24	0.23	
New Uwi	0.24	0.36	
Nganion	-0.05	-0.01	
Ngoque Blicri	-0.02	-0.07	
Nkasland Cronestan	0.41	-0.05	
People's Land of Maneau	0.09	-0.02	
Quewenia	-0.08	-0.08	
Sobianitedrucy	0.05	0.01	
Southern Ristan	-0.10	0.07	
Varijitri Isles	0.38	-0.10	
Xikong	-0.01	0.05	

	Penalty Kicks PKsv	Penalty Kicks PKm	\
Nation			
Bernepamar	0.100	-0.030	
Byasier Pujan	0.050	-0.080	
Djipines	-0.060	0.010	
Eastern Niasland	0.360	0.030	
Eastern Sleboubé	-0.030	0.035	
Esia	0.060	0.050	
Galamily	0.030	0.010	
Giumle Lizeibon	0.010	-0.040	
Greri Landmoslands	0.060	0.010	
Ledian	0.290	0.050	
Leoneku Guidisia	0.310	0.070	
Manlisgamncent	-0.070	-0.070	
Mico	0.170	-0.020	
New Uwi	0.020	-0.100	
Nganion	0.000	0.000	
Ngoque Blicri	-0.010	0.040	

Nkasland Cronestan	-0.080	0.400
People's Land of Maneau	0.060	-0.080
Quewenia	0.000	-0.010
Sobianitedrucy	-0.005	-0.020
Southern Ristan	0.040	-0.040
Varijitri Isles	0.000	0.350
Xikong	-0.040	0.000

	Number Players	TOP10	Score
Nation			
Bernepamar	21	1	8
Byasier Pujan	22	0	15
Djipines	18	0	16
Eastern Niasland	21	0	23
Eastern Sleboube	20	0	19
Esia	19	0	14
Galamily	21	1	7
Giumle Lizeibon	22	1	10
Greri Landmoslands	19	0	11
Ledian	20	0	18
Leoneku Guidisia	19	0	17
Manlisgamncent	21	0	13
Mico	21	1	4
New Uwi	18	0	20
Nganion	21	1	3
Ngoque Blicri	20	0	21
Nkasland Cronestan	19	0	22
People's Land of Maneau	21	1	2
Quewenia	20	1	5
Sobianitedrucy	25	1	1
Southern Ristan	24	1	6
Varijitri Isles	18	0	24
Xikong	19	0	12

[23 rows x 88 columns]

```
[269]: cond_filter = df.isna().sum() == 0
df = df.loc[:,cond_filter]
```

```
[270]: df
```

```
[270]:
```

	Age	Born	90s_x	Gls \
Nation				
Bernepamar	25.952381	1994.047619	2.776667	0.298571
Byasier Pujan	26.727273	1993.409091	2.086667	0.104286
Djipines	24.944444	1995.166667	2.435000	0.043889
Eastern Niasland	26.857143	1993.190476	1.594762	0.021905

Eastern Sleboube	27.100000	1993.000000	1.637500	0.051000
Esia	26.789474	1993.421053	2.301579	0.103158
Galamily	26.714286	1993.428571	2.587143	0.091905
Giumle Lizeibon	27.136364	1992.909091	2.175455	0.141364
Greri Landmoslands	27.368421	1992.631579	2.244211	0.085263
Ledian	27.800000	1992.300000	1.644000	-0.000500
Leoneku Guidisia	27.421053	1992.578947	1.748947	0.012632
Manlisgamncent	26.857143	1993.285714	2.263810	0.220000
Mico	26.285714	1993.761905	3.291429	0.182857
New Uwi	27.666667	1992.555556	1.831111	0.037778
Nganion	25.523810	1994.714286	3.684286	0.127619
Ngoque Blicri	25.700000	1994.300000	1.631500	0.018500
Nkasland Cronestan	26.052632	1994.105263	1.715789	0.116316
People's Land of Maneau	24.333333	1995.761905	4.027619	0.090000
Quewenia	27.000000	1993.050000	3.097000	0.111500
Sobianitedrucy	26.840000	1993.240000	3.510400	0.135600
Southern Ristan	28.125000	1991.958333	2.299167	0.089167
Varijitri Isles	24.444444	1995.611111	1.849444	0.024444
Xikong	28.315789	1991.789474	2.463684	0.105263

	Standard Sh	Standard SoT	Standard SoT%	\
Nation				
Bernepamar	1.092857	0.516190	36.777500	
Byasier Pujan	1.053810	0.251905	28.472143	
Djipines	0.690000	0.180556	24.854167	
Eastern Niasland	1.351905	0.166667	23.220000	
Eastern Sleboube	0.580000	0.098500	27.222222	
Esia	1.135789	0.318947	35.269286	
Galamily	1.230000	0.347619	22.672500	
Giumle Lizeibon	1.825909	0.963182	30.498000	
Greri Landmoslands	1.592105	0.433684	43.400833	
Ledian	2.155500	0.027500	10.913636	
Leoneku Guidisia	0.744211	0.068947	17.377500	
Manlisgamncent	1.197619	0.386667	31.392000	
Mico	1.614762	0.589048	38.660000	
New Uwi	0.428889	0.203889	49.580000	
Nganion	1.818095	0.632857	32.285000	
Ngoque Blicri	1.994000	0.303500	24.137333	
Nkasland Cronestan	1.714211	0.304737	21.028462	
People's Land of Maneau	0.627143	0.222857	37.751538	
Quewenia	1.217500	0.402500	24.178125	
Sobianitedrucy	1.518800	0.400400	29.788947	
Southern Ristan	0.726250	0.261667	29.703125	
Varijitri Isles	1.265556	0.371667	26.496429	
Xikong	0.786316	0.211579	28.376154	

Standard Sh/90   Standard SoT/90   Standard G/Sh   ...   \

Nation				...
Bernepamar	1.226667	0.571905	1.441667e-01	...
Byasier Pujan	1.034286	0.244762	6.428571e-02	...
Djipines	0.704444	0.170556	8.583333e-02	...
Eastern Niasland	1.266667	0.146190	1.000000e-01	...
Eastern Sleboubé	0.583000	0.099500	3.111111e-02	...
Esia	1.111579	0.310526	4.571429e-02	...
Galamily	1.268095	0.340952	2.937500e-02	...
Giumle Lizeibon	2.025909	1.146364	6.666667e-02	...
Greri Landmoslands	1.591579	0.485789	7.083333e-02	...
Ledian	2.563500	0.014500	6.545455e-02	...
Leoneku Guidisia	0.713684	0.053684	-3.469447e-18	...
Manlisgamncent	1.214762	0.374762	1.246667e-01	...
Mico	1.660476	0.630952	9.812500e-02	...
New Uwi	0.418889	0.201667	2.133333e-01	...
Nganion	1.826190	0.641429	5.888889e-02	...
Ngoque Blicri	1.897500	0.308000	4.200000e-02	...
Nkasland Cronestan	1.524211	0.327368	5.076923e-02	...
People's Land of Maneau	0.590476	0.226667	1.715385e-01	...
Quewenia	1.223500	0.388500	1.068750e-01	...
Sobianitedrucy	1.542800	0.418400	1.010526e-01	...
Southern Ristan	0.712917	0.275000	6.500000e-02	...
Varijitri Isles	1.260556	0.356667	4.571429e-02	...
Xikong	0.783158	0.212105	3.692308e-02	...

	L	Performance CS	Performance CS% \
Nation			
Bernepamar	0.530	0.080	0.020
Byasier Pujan	0.270	0.580	50.000
Djipines	0.550	0.300	25.070
Eastern Niasland	0.990	-0.050	-0.080
Eastern Sleboubé	0.805	0.230	25.045
Esia	0.540	0.040	-0.040
Galamily	0.500	0.300	40.040
Giumle Lizeibon	0.550	0.270	24.930
Greri Landmoslands	0.480	0.290	24.930
Ledian	0.630	-0.020	-0.040
Leoneku Guidisia	0.640	0.270	33.330
Manlisgamncent	0.540	0.010	0.100
Mico	0.540	0.070	16.740
New Uwi	0.270	-0.010	-0.100
Nganion	-0.050	0.290	33.380
Ngoque Blicri	0.710	0.410	33.310
Nkasland Cronestan	0.670	-0.040	-0.040
People's Land of Maneau	-0.070	0.580	71.400
Quewenia	0.260	-0.060	0.000
Sobianitedrucy	-0.020	0.645	42.960



Southern Ristan	0.220	0.620	59.920
Varijitri Isles	0.960	-0.090	-0.020
Xikong	0.320	0.440	50.090

	Performance PKatt_y	Penalty Kicks PKA \
Nation		
Bernepamar	0.13	0.08
Byasier Pujan	-0.04	0.09
Djipines	0.07	0.07
Eastern Niasland	0.41	-0.10
Eastern Sleboubé	-0.09	0.00
Esia	-0.09	-0.04
Galamily	-0.04	0.04
Giumle Lizeibon	0.31	0.16
Greri Landmoslands	0.21	0.22
Ledian	0.68	0.39
Leoneku Guidisia	0.28	-0.05
Manlisgamncent	0.23	0.27
Mico	0.24	0.23
New Uwi	0.24	0.36
Nganion	-0.05	-0.01
Ngoque Blicri	-0.02	-0.07
Nkasland Cronestan	0.41	-0.05
People's Land of Maneau	0.09	-0.02
Quewenia	-0.08	-0.08
Sobianitedrucy	0.05	0.01
Southern Ristan	-0.10	0.07
Varijitri Isles	0.38	-0.10
Xikong	-0.01	0.05

	Penalty Kicks PKsv	Penalty Kicks PKm \
Nation		
Bernepamar	0.100	-0.030
Byasier Pujan	0.050	-0.080
Djipines	-0.060	0.010
Eastern Niasland	0.360	0.030
Eastern Sleboubé	-0.030	0.035
Esia	0.060	0.050
Galamily	0.030	0.010
Giumle Lizeibon	0.010	-0.040
Greri Landmoslands	0.060	0.010
Ledian	0.290	0.050
Leoneku Guidisia	0.310	0.070
Manlisgamncent	-0.070	-0.070
Mico	0.170	-0.020
New Uwi	0.020	-0.100
Nganion	0.000	0.000

Ngoque Blicri	-0.010	0.040
Nkasland Cronestan	-0.080	0.400
People's Land of Maneau	0.060	-0.080
Quewenia	0.000	-0.010
Sobianitedrucy	-0.005	-0.020
Southern Ristan	0.040	-0.040
Varijitri Isles	0.000	0.350
Xikong	-0.040	0.000

	Number Players	TOP10	Score
Nation			
Bernepamar	21	1	8
Byasier Pujan	22	0	15
Djipines	18	0	16
Eastern Niasland	21	0	23
Eastern Sleboubé	20	0	19
Esia	19	0	14
Galamily	21	1	7
Giumle Lizeibon	22	1	10
Greri Landmoslands	19	0	11
Ledian	20	0	18
Leoneku Guidisia	19	0	17
Manlisgamncent	21	0	13
Mico	21	1	4
New Uwi	18	0	20
Nganion	21	1	3
Ngoque Blicri	20	0	21
Nkasland Cronestan	19	0	22
People's Land of Maneau	21	1	2
Quewenia	20	1	5
Sobianitedrucy	25	1	1
Southern Ristan	24	1	6
Varijitri Isles	18	0	24
Xikong	19	0	12

[23 rows x 88 columns]

```
[271]: # LogReg
y = df.TOP10.values
X = df.drop(["TOP10", "Score"], axis=1)

logreg = LogisticRegression(max_iter=1000, penalty="l1", solver='liblinear')
logreg.fit(X.values, y)

display(pd.DataFrame({"Nation": df.index, "Score": df.Score, "Prediction": logreg.
    ↪ predict(X)}))
```

```
logreg_coefs = pd.DataFrame({"Stat":X.columns, "Coef":logreg.coef_[0]})
display(logreg_coefs[logreg_coefs.Coef != 0])
```

	Nation	Score	Prediction
Nation			
Bernepamar	Bernepamar	8	1
Byasier Pujan	Byasier Pujan	15	0
Djipines	Djipines	16	0
Eastern Niasland	Eastern Niasland	23	0
Eastern Sleboube	Eastern Sleboube	19	0
Esia	Esia	14	0
Galamily	Galamily	7	1
Giumle Lizeibon	Giumle Lizeibon	10	1
Greri Landmoslands	Greri Landmoslands	11	0
Ledian	Ledian	18	0
Leoneku Guidisia	Leoneku Guidisia	17	0
Manlisgamncent	Manlisgamncent	13	0
Mico	Mico	4	1
New Uwi	New Uwi	20	0
Nganion	Nganion	3	1
Ngoque Blicri	Ngoque Blicri	21	0
Nkasland Cronestan	Nkasland Cronestan	22	0
People's Land of Maneau	People's Land of Maneau	2	1
Quewenia	Quewenia	5	1
Sobianitedrucy	Sobianitedrucy	1	1
Southern Ristan	Southern Ristan	6	1
Varijitri Isles	Varijitri Isles	24	0
Xikong	Xikong	12	0

	Stat	Coef
1	Born	-0.052235
24	Total TotDist	-0.012942
25	Total PrgDist	0.110752
34	Long Cmp%	0.318370
69	Playing Time Min	0.118015
75	Performance Save%	0.371513

```
[272]: # Lasso regression
y = df.Score.values
X = df.drop(["TOP10", "Score"], axis=1)
X = X.drop(["Playing Time Min", "90s", "90s_x", "90s_y",
            "Playing Time 90s", "W", "L", "Playing Time MP",
            "Playing Time Starts", "Performance PKatt_x"], axis=1)
# X = X[["Born", "Standard SoT%", "Total TotDist", "Total PrgDist", "Short Cmp%"]]
X = X[["Gls", "Standard FK", "Total Cmp%", "Medium Cmp%", "Blocks_
↪Pass", "Performance GA",
        "Penalty Kicks PKm", "Penalty Kicks PKm"]]
X = (X-X.mean())/X.std()
```

```

lasso = Lasso(tol=0.01, alpha=1)
lasso.fit(X.values, y)

lasso_coefs = pd.DataFrame({"Stat":X.columns, "Coef":lasso.coef_})
lasso_coefs = lasso_coefs[lasso_coefs.Coeff != 0]
display(lasso_coefs)

print("R2 score", r2_score(y, lasso.predict(X)))
print("Cross Validation Scores", cross_val_score(lasso, X, y))
print("Mean Cross Validation Score", np.mean(cross_val_score(lasso, X, y)))

display(pd.DataFrame({"Score":df.Score, "Prediction":lasso.predict(X)}))

```

	Stat	Coef
0	Gls	-1.435584e+00
1	Standard FK	-9.250601e-01
2	Total Cmp%	-8.746235e-02
3	Medium Cmp%	-4.047183e-01
4	Blocks Pass	3.061693e-01
5	Performance GA	3.683531e+00
6	Penalty Kicks PKm	4.936185e-01
7	Penalty Kicks PKm	9.689219e-16

R2 score 0.8108638962366675

Cross Validation Scores [0.5582736 0.70296792 0.66002075 0.56450478 0.9117763 ]

Mean Cross Validation Score 0.6795086684120664

	Score	Prediction
Nation		
Bernepamar	8	9.594306
Byasier Pujan	15	10.053262
Djipines	16	15.913205
Eastern Niasland	23	20.171063
Eastern Sleboube	19	19.215044
Esia	14	12.177759
Galamily	7	8.853006
Giumle Lizeibon	10	10.288218
Greri Landmoslands	11	12.812316
Ledian	18	19.524193
Leoneku Guidisia	17	13.561877
Manlisgamncent	13	11.637444
Mico	4	8.263474
New Uwi	20	17.416598
Nganion	3	5.547991
Ngoque Blicri	21	16.015472
Nkasland Cronestan	22	18.064685
People's Land of Maneau	2	5.734234
Qewenia	5	12.486205

Sobianitedrucy	1	2.502652
Southern Ristan	6	8.085331
Varijitri Isles	24	21.713285
Xikong	12	11.368380

```
[369]: # Lasso regression
y = df.Score.values

X = df.drop(["TOP10", "Score"], axis=1)
X = X.drop(["Playing Time Min", "90s", "90s_x", "90s_y"], axis=1)
# X = X[["Total TotDist", "Performance CS%", "Standard SoT%"]]
# X = X.drop(["Born", "Standard SoT%", "Total TotDist", "Total PrgDist", "Short_
↳ Cmp%"], axis=1)
# X = X[["Born", "Standard SoT%", "Total TotDist", "Total PrgDist", "Short_
↳ Cmp%", "90s"]]
Y = X
X = (X - X.mean()) / X.std()
X = X[["Gls", "Standard FK", "Total Cmp%", "Medium Cmp%", "Blocks_
↳ Pass", "Performance GA",
      "Penalty Kicks PKm"]]
# X = X.drop("Number Players", axis=1)

lasso = LassoCV(cv=3, tol=0.1)
lasso.fit(X.values, y)

lasso_coefs = pd.DataFrame({"Stat": X.columns, "Coef": lasso.coef_})
lasso_coefs = lasso_coefs[lasso_coefs.Coef != 0]
display(lasso_coefs)

y_pred = lasso.predict(X)

print(r2_score(y, y_pred))
print(np.mean(cross_val_score(lasso, X, y)))

display(pd.DataFrame({"Score": df.Score, "Prediction": lasso.predict(X)}))
```

	Stat	Coef
0	Gls	-1.606882
1	Standard FK	-0.948223
2	Total Cmp%	-0.397181
3	Medium Cmp%	-0.896960
4	Blocks Pass	0.909493
5	Performance GA	4.089069
6	Penalty Kicks PKm	0.933866

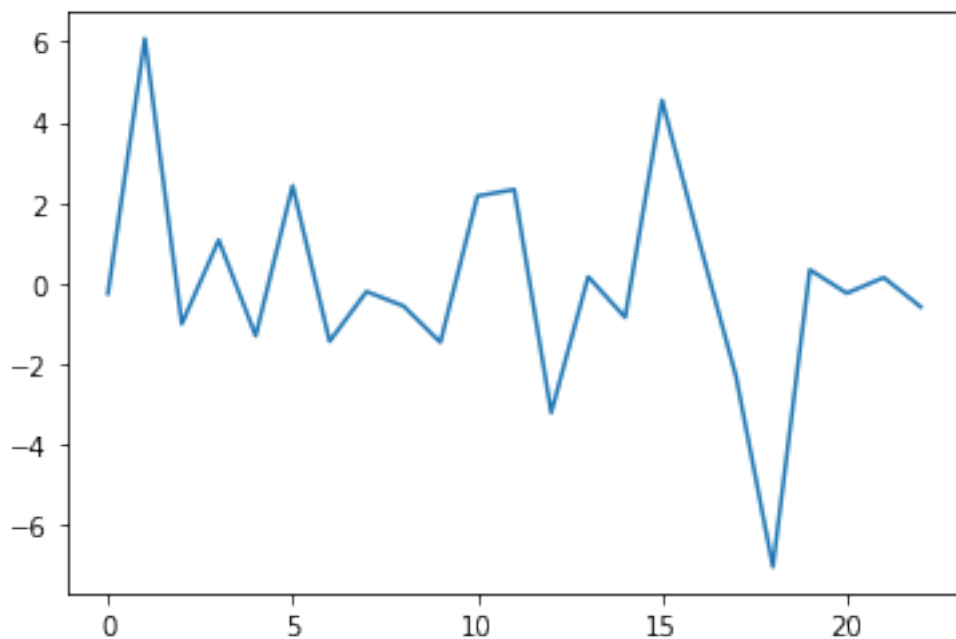
0.8686162614153354  
0.7788877250022239

Score Prediction

Nation		
Bernepamar	8	8.251070
Byasier Pujan	15	8.919039
Djipines	16	17.000062
Eastern Niasland	23	21.919632
Eastern Sleboube	19	20.295710
Esia	14	11.572838
Galamily	7	8.432117
Giumle Lizeibon	10	10.198040
Greri Landmoslands	11	11.555928
Ledian	18	19.456228
Leoneku Guidisia	17	14.829796
Manlisgamncent	13	10.666611
Mico	4	7.201661
New Uwi	20	19.830021
Nganion	3	3.836707
Ngoque Blicri	21	16.452556
Nkasland Cronestan	22	20.938284
People's Land of Maneau	2	4.294670
Quewenia	5	12.026963
Sobianitedrucus	1	0.656990
Southern Ristan	6	6.237996
Varijitri Isles	24	23.850353
Xikong	12	12.576727

```
[425]: plt.plot(y - y_pred)
```

```
[425]: [<matplotlib.lines.Line2D at 0x7fd5d28a36a0>]
```



```
[426]: # np.mean(abs(y-y_pred))
np.var(y-y_pred)
```

```
[426]: 6.496168367524549
```

```
[427]: np.std(y-y_pred)
```

```
[427]: 2.548758201070582
```

```
[274]: contributions_df = X[lasso_coefs.Stat.values]
for stat in contributions_df.columns:
    contributions_df[stat] *= lasso_coefs.loc[lasso_coefs.Stat == stat, "Coef"].
    ↪ values
print("Intercept is {}".format(lasso.intercept_))
display(contributions_df)
display(contributions_df.var())
```

```
Intercept is 12.652173913043473
```

	Gls	Standard FK	Total Cmp%	Medium Cmp%	\
Nation					
Bernepamar	-4.585122	-0.496384	-0.394654	-1.036835	
Byasier Pujan	-0.183277	0.591883	-0.008756	0.132088	
Djipines	1.185107	1.361100	0.058592	-0.120066	
Eastern Niasland	1.683191	-0.391491	-0.069817	0.037137	
Eastern Sleboube	1.023993	-0.145647	0.376373	0.245497	
Esia	-0.157725	-1.149206	-0.510880	-1.043028	
Galamily	0.097232	0.146087	0.429727	0.276089	
Giumle Lizeibon	-1.023335	0.649097	-0.339027	-0.656804	
Greri Landmoslands	0.247708	-0.526058	-0.445167	-1.161921	
Ledian	2.190806	0.941964	-0.423702	-0.949153	
Leoneku Guidisia	1.893289	1.126010	0.824112	0.340044	
Manlisgamncent	-2.804964	-0.194816	-0.076459	-0.071488	
Mico	-1.963435	-0.443937	-0.069276	0.016524	
New Uwi	1.323563	0.473878	0.418312	1.116492	
Nganion	-0.711930	-2.148451	-0.499615	-1.140264	
Ngoque Blicri	1.760331	0.762990	0.262020	0.465856	
Nkasland Cronestan	-0.455837	0.676764	0.543750	1.787329	
People's Land of Maneau	0.140388	-0.339044	-0.249495	0.112007	
Quewenia	-0.346728	0.267370	0.328416	0.078440	
Sobianitedrucy	-0.892751	-2.466802	-0.311074	0.053154	
Southern Ristan	0.159268	0.370624	-0.389019	-0.574561	
Varijitri Isles	1.625651	0.489175	-0.027738	-0.378498	
Xikong	-0.205423	0.444895	0.573376	2.471958	

```
Blocks Pass Performance GA Penalty Kicks PKm
```

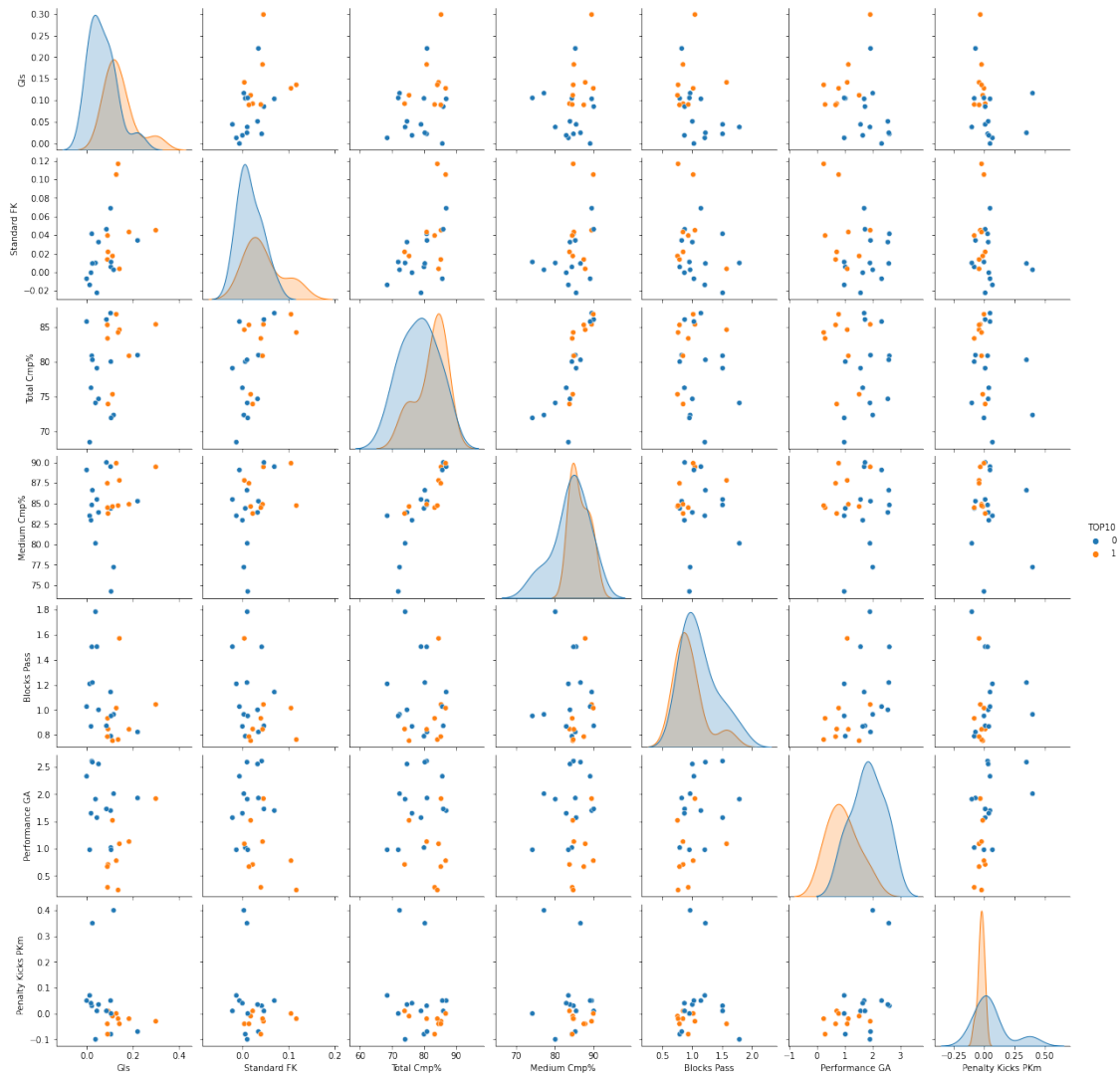
Nation			
Bernepamar	-0.029355	2.566929	-0.425684
Byasier Pujan	-0.829774	-2.619545	-0.815753
Djipines	1.426818	0.549967	-0.113629
Eastern Niasland	1.422814	6.543225	0.042399
Eastern Sleboubé	-0.164360	6.226274	0.081406
Esia	0.283953	1.299124	0.198426
Galamily	-0.649567	-4.405997	-0.113629
Giumle Lizeibon	1.635786	-2.216153	-0.503698
Greri Landmoslands	-0.569185	1.472006	-0.113629
Ledian	-0.083942	4.929655	0.198426
Leoneku Guidisia	0.489768	-2.850055	0.354454
Manlisgamncent	-0.724653	2.624556	-0.737739
Mico	-0.657076	-1.985643	-0.347670
New Uwi	2.308081	2.509301	-0.971780
Nganion	-0.120960	-4.002605	-0.191643
Ngoque Blicri	-0.582215	1.010986	0.120413
Nkasland Cronestan	-0.280380	3.085576	2.928908
People's Land of Maneau	-0.379257	-6.826351	-0.815753
Quewenia	-0.944881	0.261829	-0.269656
Sobianitedrucy	-0.915553	-7.114489	-0.347670
Southern Ristan	-0.840286	-4.636507	-0.503698
Varijitri Isles	0.522780	6.427970	2.538839
Xikong	-0.318555	-2.850055	-0.191643
Gls	2.582070		
Standard FK	0.899127		
Total Cmp%	0.157752		
Medium Cmp%	0.804538		
Blocks Pass	0.827178		
Performance GA	16.720482		
Penalty Kicks PKm	0.872105		
dtype: float64			

```
[275]: # useful_cols = ["Born", "Total TotDist", "Total PrgDist", "Medium Cmp%", "Long
↳ Cmp%",
#           "Pressures Press", "Playing Time Min", "Performance SoTA",
↳ "Performance Save%", "TOP10"]
useful_cols = list(contributions_df.columns) + ["TOP10"]
Test = df[useful_cols]
```

```
[276]: sns.pairplot(Test, hue="TOP10")
```

```
[276]: <seaborn.axisgrid.PairGrid at 0x7fd5c9774550>
```





```
[350]: rarita["Player"]
```

```
[350]: 0      H. Tourgeman
      1      M. Ludwig
      2      E. Nakanjako
      3      H. Mubaiwa
      4      E. Koç
      ...
     168      H. Amade
     169      H. Mireembe
     170      Y. Acola
     171      Y. Draru
     172      W. Nasiru
     Name: Player, Length: 173, dtype: object
```

[353]: rarita

```
[353]:
```

	Player	Pos	1/3	90s	A-xA	Age	Ast	\
0	H. Tourgeman	MFFW	-0.405518	-0.609547	0.022125	0.231471	-0.284255	
1	M. Ludwig	DFMF	-0.477353	1.386093	-0.076422	1.227602	-0.351675	
2	E. Nakanjako	MFFW	-0.463772	0.251861	0.137095	-1.163112	-0.174698	
3	H. Mubaiwa	FW	-0.549189	-1.162351	-0.207817	1.227602	-0.360103	
4	E. Koç	DF	-0.490577	0.050597	-0.027149	0.430697	-0.309538	
..	...	...	...	...	...	...	...	
168	H. Amade	MFFW	-0.423387	-0.764296	-0.224241	-1.163112	-0.469661	
169	H. Mirembe	MFFW	-0.547044	-0.996868	0.531281	-1.362339	-0.006147	
170	Y. Acola	GK	NaN	NaN	NaN	1.426828	NaN	
171	Y. Draru	GK	NaN	NaN	NaN	2.223733	NaN	
172	W. Nasiru	GK	NaN	NaN	NaN	1.426828	NaN	

	Blocks	Blocks	Blocks	Pass	Blocks	Sh	...	Total	PrgDist	Total	TotDist	\
0	1.070342		1.459323		-0.620752		...	-0.466709		-0.509406		
1	0.346984		0.023462		0.642900		...	-0.482412		-0.545518		
2	-0.328149		-0.022121		-0.775485		...	-0.501488		-0.541997		
3	-1.340849		-1.241464		-0.698119		...	-0.513126		-0.549453		
4	1.504356		1.812591		0.256068		...	-0.430759		-0.515640		
..	...		...		...	...	...	...		...		
168	-0.125609		0.342542		-0.594963		...	-0.486663		-0.528094		
169	0.009418		-0.079100		0.591322		...	-0.521754		-0.577631		
170	NaN		NaN		NaN		...	NaN		NaN		
171	NaN		NaN		NaN		...	NaN		NaN		
172	NaN		NaN		NaN		...	NaN		NaN		

	Vs	Dribbles	Att	Vs	Dribbles	Past	Vs	Dribbles	Tkl	Vs	Dribbles	Tkl%	\
0	0.385442			0.526248			-0.420673			-0.780674			
1	0.594336			0.005009			1.475841			0.887800			
2	-0.443394			-0.340036			-0.203364			0.479901			
3	-1.009428			-0.839250			-0.835535			NaN			
4	0.594336			0.386761			0.547339			-0.024418			
..	...			...			...			...			
168	0.567382			0.511565			0.231254			-0.251473			
169	-0.719672			-0.523570			-1.151621			-1.468059			
170	NaN			NaN			NaN			NaN			
171	NaN			NaN			NaN			NaN			
172	NaN			NaN			NaN			NaN			

	W	Year	xA	Value
0	NaN	NaN	-0.390782	-3.024274
1	NaN	NaN	-0.572416	1.372657
2	NaN	NaN	-0.299964	-1.890153
3	NaN	NaN	-0.561064	1.426557
4	NaN	NaN	-0.379430	3.266117

```

..      ...      ...      ...      ...
168      NaN      NaN      -0.458895      -1.663179
169      NaN      NaN      -0.504303      1.099757
170      0.523721      NaN      NaN      0.978324
171      0.391611      NaN      NaN      -0.692210
172      0.589776      NaN      NaN      -2.458530

```

[173 rows x 89 columns]

```

[358]: not_normalised = leagues_merged_df[leagues_merged_df["Nation"] == "Rarita"]
not_normalised = not_normalised.reset_index(drop=True)
not_normalised["Value"] = rarita["Value"]
not_normalised.to_csv("not_normalised.csv")

```

```

[360]: not_normalised

```

```

[360]:      Player  Nation  Pos      Squad_x  Age  Born  90s_x  \
0  H. Tourgeman  Rarita  MFFW      Fanatical Outlaws  27  1993   6.22
1    M. Ludwig  Rarita  DFMF  Great Galactic Gorgons  32  1988  28.56
2  E. Nakanjako  Rarita  MFFW      Mighty Jays  20  2000  16.03
3    H. Mubaiwa  Rarita  FW      Solemn Cougars  32  1988   0.27
4    E. Koç  Rarita  DF      Stacked Rocks  28  1992  13.88
..      ...      ...      ...      ...      ...      ...
168    H. Amade  Rarita  MFFW      Wild Hornets  20  2000   4.69
169    H. Mirembe  Rarita  MFFW      Wild Hornets  19  2001   2.18
170    Y. Acola  Rarita  GK      NaN  33  1987   NaN
171    Y. Draru  Rarita  GK      NaN  37  1983   NaN
172    W. Nasiru  Rarita  GK      NaN  33  1986   NaN

```

```

      GlS  Standard Sh  Standard SoT  ...  D  L  Performance CS  \
0  0.26      1.91      0.53  ...  NaN  NaN      NaN
1  0.01      0.36      0.06  ...  NaN  NaN      NaN
2  0.14      0.74      0.23  ...  NaN  NaN      NaN
3  0.06      5.07     -0.06  ...  NaN  NaN      NaN
4  0.08      0.18      0.16  ...  NaN  NaN      NaN
..      ...      ...      ...      ...      ...
168  0.25      1.94      0.64  ...  NaN  NaN      NaN
169  0.09      1.83     -0.04  ...  NaN  NaN      NaN
170  NaN      NaN      NaN  ...  0.34  0.23   0.16
171  NaN      NaN      NaN  ...  0.31  0.27   0.29
172  NaN      NaN      NaN  ...  0.20  0.27   0.33

```

```

      Performance CS%  Performance PKatt_y  Penalty Kicks PKA  \
0      NaN      NaN      NaN
1      NaN      NaN      NaN
2      NaN      NaN      NaN
3      NaN      NaN      NaN

```

4	NaN	NaN	NaN
..	...	...	...
168	NaN	NaN	NaN
169	NaN	NaN	NaN
170	18.80	0.08	0.13
171	34.92	0.30	0.27
172	31.68	0.19	0.22

	Penalty Kicks	PKsv	Penalty Kicks	PKm	Penalty Kicks	Save%	Value
0		NaN		NaN		NaN	-3.024274
1		NaN		NaN		NaN	1.372657
2		NaN		NaN		NaN	-1.890153
3		NaN		NaN		NaN	1.426557
4		NaN		NaN		NaN	3.266117
..		...		...		...	...
168		NaN		NaN		NaN	-1.663179
169		NaN		NaN		NaN	1.099757
170		-0.05		0.07		19.92	0.978324
171		-0.03		0.08		20.01	-0.692210
172		0.07		0.04		0.06	-2.458530

[173 rows x 97 columns]

```
[329]: rarita = leagues_merged_df[leagues_merged_df.Nation == "Rarita"]
# rarita = leagues_merged_df[leagues_merged_df.League == "RFL"]

rarita = rarita.drop(["Squad_x", "Squad_y", "League"], axis=1)
rarita
```

```
[329]:
```

	Player	Nation	Pos	Age	Born	90s_x	Gls	Standard	Sh	\
23	H. Tourgeman	Rarita	MFFW	27	1993	6.22	0.26		1.91	
77	M. Ludwig	Rarita	DFMF	32	1988	28.56	0.01		0.36	
213	E. Nakanjako	Rarita	MFFW	20	2000	16.03	0.14		0.74	
315	H. Mubaiwa	Rarita	FW	32	1988	0.27	0.06		5.07	
361	E. Koç	Rarita	DF	28	1992	13.88	0.08		0.18	
...	...	...	...	...	...	...	...		...	
3016	H. Amade	Rarita	MFFW	20	2000	4.69	0.25		1.94	
3017	H. Mirembe	Rarita	MFFW	19	2001	2.18	0.09		1.83	
3018	Y. Acola	Rarita	GK	33	1987	NaN	NaN		NaN	
3019	Y. Draru	Rarita	GK	37	1983	NaN	NaN		NaN	
3020	W. Nasiru	Rarita	GK	33	1986	NaN	NaN		NaN	

	Standard	SoT	Standard	SoT%	...	W	D	L	Performance	CS	\
23		0.53		24.91	...	NaN	NaN	NaN		NaN	
77		0.06		10.05	...	NaN	NaN	NaN		NaN	
213		0.23		30.88	...	NaN	NaN	NaN		NaN	
315		-0.06		-0.02	...	NaN	NaN	NaN		NaN	

361	0.16	33.38	...	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...
3016	0.64	33.34	...	NaN	NaN	NaN	NaN
3017	-0.04	-0.04	...	NaN	NaN	NaN	NaN
3018	NaN	NaN	...	0.51	0.34	0.23	0.16
3019	NaN	NaN	...	0.47	0.31	0.27	0.29
3020	NaN	NaN	...	0.53	0.20	0.27	0.33

	Performance	CS%	Performance	PKatt_y	Penalty Kicks	PKA	\
23		NaN		NaN		NaN	
77		NaN		NaN		NaN	
213		NaN		NaN		NaN	
315		NaN		NaN		NaN	
361		NaN		NaN		NaN	
...		...		...		...	
3016		NaN		NaN		NaN	
3017		NaN		NaN		NaN	
3018		18.80		0.08		0.13	
3019		34.92		0.30		0.27	
3020		31.68		0.19		0.22	

	Penalty Kicks	PKsv	Penalty Kicks	PKm	Penalty Kicks	Save%
23		NaN		NaN		NaN
77		NaN		NaN		NaN
213		NaN		NaN		NaN
315		NaN		NaN		NaN
361		NaN		NaN		NaN
...		...		...		...
3016		NaN		NaN		NaN
3017		NaN		NaN		NaN
3018		-0.05		0.07		19.92
3019		-0.03		0.08		20.01
3020		0.07		0.04		0.06

[173 rows x 91 columns]

```
[330]: rarita = rarita.set_index(["Player", "Pos"])
```

```
[331]: Y = Y[lasso_coefs.Stat.values]
```

```
[332]: standard_rarita = (rarita - rarita.mean())/rarita.std()

# standard_rarita = rarita[lasso_coefs.Stat.values]
standard_rarita = (standard_rarita - standard_rarita.mean()) / standard_rarita.
    ↪std()
standard_rarita
```

[332]:

		1/3	90s	90s_x	90s_y	A-xA	Age	\
Player	Pos							
H. Tourgeman	MFFW	-0.405518	-0.609547	-0.619892	-0.615964	0.022125	0.231471	
M. Ludwig	DFMF	-0.477353	1.386093	1.378076	1.374600	-0.076422	1.227602	
E. Nakanjako	MFFW	-0.463772	0.251861	0.257461	0.263962	0.137095	-1.163112	
H. Mubaiwa	FW	-0.549189	-1.162351	-1.152027	-1.163235	-0.207817	1.227602	
E. Koç	DF	-0.490577	0.050597	0.065177	0.066336	-0.027149	0.430697	
...	...	...	...	...	...	...	...	
H. Amade	MFFW	-0.423387	-0.764296	-0.756727	-0.772455	-0.224241	-1.163112	
H. Mirembe	MFFW	-0.547044	-0.996868	-0.981207	-0.995119	0.531281	-1.362339	
Y. Acola	GK	NaN	NaN	NaN	NaN	NaN	1.426828	
Y. Draru	GK	NaN	NaN	NaN	NaN	NaN	2.223733	
W. Nasiru	GK	NaN	NaN	NaN	NaN	NaN	1.426828	

		Ast	Blocks	Blocks	Blocks	Pass	Blocks	Sh	...	\
Player	Pos									
H. Tourgeman	MFFW	-0.284255		1.070342		1.459323	-0.620752		...	
M. Ludwig	DFMF	-0.351675		0.346984		0.023462	0.642900		...	
E. Nakanjako	MFFW	-0.174698		-0.328149		-0.022121	-0.775485		...	
H. Mubaiwa	FW	-0.360103		-1.340849		-1.241464	-0.698119		...	
E. Koç	DF	-0.309538		1.504356		1.812591	0.256068		...	
...	...	...		...		...	...		...	
H. Amade	MFFW	-0.469661		-0.125609		0.342542	-0.594963		...	
H. Mirembe	MFFW	-0.006147		0.009418		-0.079100	0.591322		...	
Y. Acola	GK	NaN		NaN		NaN	NaN		...	
Y. Draru	GK	NaN		NaN		NaN	NaN		...	
W. Nasiru	GK	NaN		NaN		NaN	NaN		...	

		Total Cmp%	Total PrgDist	Total TotDist	Vs Dribbles	Att	\
Player	Pos						
H. Tourgeman	MFFW	1.068104	-0.466709	-0.509406		0.385442	
M. Ludwig	DFMF	0.390836	-0.482412	-0.545518		0.594336	
E. Nakanjako	MFFW	0.964746	-0.501488	-0.541997		-0.443394	
H. Mubaiwa	FW	-0.396137	-0.513126	-0.549453		-1.009428	
E. Koç	DF	0.324650	-0.430759	-0.515640		0.594336	
...	...	...	...	...		...	
H. Amade	MFFW	0.340063	-0.486663	-0.528094		0.567382	
H. Mirembe	MFFW	-0.612827	-0.521754	-0.577631		-0.719672	
Y. Acola	GK	NaN	NaN	NaN		NaN	
Y. Draru	GK	NaN	NaN	NaN		NaN	
W. Nasiru	GK	NaN	NaN	NaN		NaN	

		Vs Dribbles Past	Vs Dribbles Tkl	Vs Dribbles Tkl%	\
Player	Pos				
H. Tourgeman	MFFW	0.526248	-0.420673	-0.780674	
M. Ludwig	DFMF	0.005009	1.475841	0.887800	
E. Nakanjako	MFFW	-0.340036	-0.203364	0.479901	

H. Mubaiwa	FW	-0.839250	-0.835535	NaN
E. Koç	DF	0.386761	0.547339	-0.024418
...		...	...	...
H. Amade	MFFW	0.511565	0.231254	-0.251473
H. Mirembe	MFFW	-0.523570	-1.151621	-1.468059
Y. Acola	GK	NaN	NaN	NaN
Y. Draru	GK	NaN	NaN	NaN
W. Nasiru	GK	NaN	NaN	NaN

Player	Pos	W	Year	xA
H. Tourgeman	MFFW	NaN	NaN	-0.390782
M. Ludwig	DFMF	NaN	NaN	-0.572416
E. Nakanjako	MFFW	NaN	NaN	-0.299964
H. Mubaiwa	FW	NaN	NaN	-0.561064
E. Koç	DF	NaN	NaN	-0.379430
...		...	...	...
H. Amade	MFFW	NaN	NaN	-0.458895
H. Mirembe	MFFW	NaN	NaN	-0.504303
Y. Acola	GK	0.523721	NaN	NaN
Y. Draru	GK	0.391611	NaN	NaN
W. Nasiru	GK	0.589776	NaN	NaN

[173 rows x 89 columns]

```
[334]: rarita = standard_rarita.reset_index()
rarita
```

```
[334]:
```

	Player	Pos	1/3	90s	90s_x	90s_y	A-xA	\
0	H. Tourgeman	MFFW	-0.405518	-0.609547	-0.619892	-0.615964	0.022125	
1	M. Ludwig	DFMF	-0.477353	1.386093	1.378076	1.374600	-0.076422	
2	E. Nakanjako	MFFW	-0.463772	0.251861	0.257461	0.263962	0.137095	
3	H. Mubaiwa	FW	-0.549189	-1.162351	-1.152027	-1.163235	-0.207817	
4	E. Koç	DF	-0.490577	0.050597	0.065177	0.066336	-0.027149	
..	...	...	...	...	...	...	...	
168	H. Amade	MFFW	-0.423387	-0.764296	-0.756727	-0.772455	-0.224241	
169	H. Mirembe	MFFW	-0.547044	-0.996868	-0.981207	-0.995119	0.531281	
170	Y. Acola	GK	NaN	NaN	NaN	NaN	NaN	
171	Y. Draru	GK	NaN	NaN	NaN	NaN	NaN	
172	W. Nasiru	GK	NaN	NaN	NaN	NaN	NaN	

	Age	Ast	Blocks	Blocks	...	Total Cmp%	Total PrgDist	\
0	0.231471	-0.284255	1.070342	...	1.068104	-0.466709		
1	1.227602	-0.351675	0.346984	...	0.390836	-0.482412		
2	-1.163112	-0.174698	-0.328149	...	0.964746	-0.501488		
3	1.227602	-0.360103	-1.340849	...	-0.396137	-0.513126		
4	0.430697	-0.309538	1.504356	...	0.324650	-0.430759		

```

..      ...      ...      ...      ...      ...
168 -1.163112 -0.469661      -0.125609 ...      0.340063      -0.486663
169 -1.362339 -0.006147      0.009418 ...      -0.612827      -0.521754
170  1.426828      NaN      NaN      ...      NaN      NaN
171  2.223733      NaN      NaN      ...      NaN      NaN
172  1.426828      NaN      NaN      ...      NaN      NaN

      Total TotDist  Vs Dribbles Att  Vs Dribbles Past  Vs Dribbles Tkl  \
0      -0.509406      0.385442      0.526248      -0.420673
1      -0.545518      0.594336      0.005009      1.475841
2      -0.541997      -0.443394      -0.340036      -0.203364
3      -0.549453      -1.009428      -0.839250      -0.835535
4      -0.515640      0.594336      0.386761      0.547339
..      ...      ...      ...      ...
168      -0.528094      0.567382      0.511565      0.231254
169      -0.577631      -0.719672      -0.523570      -1.151621
170      NaN      NaN      NaN      NaN
171      NaN      NaN      NaN      NaN
172      NaN      NaN      NaN      NaN

      Vs Dribbles Tkl%      W  Year      xA
0      -0.780674      NaN  NaN -0.390782
1      0.887800      NaN  NaN -0.572416
2      0.479901      NaN  NaN -0.299964
3      NaN      NaN  NaN -0.561064
4      -0.024418      NaN  NaN -0.379430
..      ...      ...      ...
168      -0.251473      NaN  NaN -0.458895
169      -1.468059      NaN  NaN -0.504303
170      NaN  0.523721      NaN      NaN
171      NaN  0.391611      NaN      NaN
172      NaN  0.589776      NaN      NaN

```

[173 rows x 91 columns]

```

[335]: df_ = rarita.loc[rarita.Pos.isin(["DF", "DFMF", "MFDF"])]
mf = rarita.loc[rarita.Pos.isin(["MF", "MFFW", "FWMF", "MFDF", "DFMF"])]
fw = rarita.loc[rarita.Pos.isin(["FW", "MFFW", "FWMF"])]
gk = rarita.loc[rarita.Pos.isin(["GK"])]

```

```

[336]: for stat, val in lasso_coefs.values:
        print(stat, val)

```

```

Gls -1.6068821243639813
Standard FK -0.9482233198338254
Total Cmp% -0.3971806653665649
Medium Cmp% -0.8969603214277939
Blocks Pass 0.909493299541201

```



Performance GA 4.089068639826528  
 Penalty Kicks PKm 0.9338657427762009

```
[337]: lasso.intercept_
```

```
[337]: 12.652173913043473
```

```
[ ]:
```

```
[347]: rarita.sort_values("Value").reset_index(drop=True).to_csv("normalised_data.csv")
```

```
[344]: rarita = rarita.drop(["90s_x", "90s_y", "Performance PKatt_x"], axis=1)
```

```
[468]: rarita["Value"] = rarita.fillna(0).apply(lambda row: sum(
    [row[stat] * coef for stat, coef in lasso_coefs.values]) , axis=1)

res = rarita.sort_values("Value", ascending=True)[
    ["Player", "Pos"] + list(lasso_coefs.Stat.values) + ["Value"]]
res = res.reset_index(drop=True)
res.head(25)
```

```
[468]:
```

	Player	Pos	Gls	Standard FK	Total Cmp%	Medium Cmp%	\
0	K. Kazlo?	FW	4.832101	3.937875	-0.578374	-0.882552	
1	B. Ampofo	GK	-0.118982	0.540325	0.321024	1.613995	
2	A. Omar	GK	-0.903910	-0.551745	0.913974	1.610083	
3	M. Asiiimwe	GK	-0.602015	-0.066381	0.185933	1.606172	
4	J. Namirembe	MF	2.115043	2.117759	-0.915648	-0.898199	
5	H. Mubaiwa	FW	2.598076	0.054960	-0.253793	-0.867884	
6	P. Rabi	MFFW	0.847083	4.059216	-0.465949	-0.853216	
7	U. Shoko	FW	2.718834	-0.551745	-0.018970	-0.073839	
8	D. Lehner	MFFW	2.477317	0.661666	-0.633680	-0.877663	
9	G. Jankowski	MF	2.537697	-0.309063	0.062628	-0.191186	
10	H. Makumbi	FW	2.175422	-0.673086	-0.062490	-0.201943	
11	P. Villa	MFFW	1.632011	-0.309063	0.884961	0.451286	
12	H. Tourgeman	MFFW	0.967841	1.753736	1.068104	0.790613	
13	L. Tarigan	FW	3.020729	-1.279792	-0.339925	-1.393010	
14	A. Kyarikunda	FW	0.967841	1.025689	0.388116	0.118805	
15	Z. Zziwa	FW	2.235801	-1.037109	0.724483	-0.102198	
16	N. Yamashita	FW	3.322624	-0.187722	-0.846743	-2.499978	
17	L. Mandala	MFFW	1.390494	-0.187722	-0.179448	0.018082	
18	T. Larsson	DF	-0.118982	0.661666	1.418978	1.180790	
19	O. Wanjala	MF	1.148978	0.054960	1.074451	0.886446	
20	B. Ayuba	MF	2.839592	0.418984	-0.094222	-1.419413	
21	W. Nasiru	GK	NaN	NaN	NaN	NaN	
22	B. Katooko	MF	-0.541636	4.180558	-0.372564	-1.259039	
23	H. Valentini	MF	-0.179361	1.753736	-0.010810	0.303625	
24	F. Akumu	GK	-0.058603	-1.401133	0.349130	1.486869	

	Blocks Pass	Performance GA	Penalty Kicks PKm	Value
0	-0.945175	NaN	NaN	-11.336899
1	-1.150298	-1.147161	-0.909313	-8.482538
2	-1.252860	-1.573163	-0.909313	-8.252954
3	-1.230068	-1.253662	-0.739575	-7.419922
4	-0.375389	NaN	NaN	-4.578821
5	-1.013549	NaN	NaN	-4.269473
6	0.194398	NaN	NaN	-4.083036
7	-0.136078	NaN	NaN	-3.895665
8	-0.250036	NaN	NaN	-3.796656
9	0.114628	NaN	NaN	-3.533854
10	-0.956571	NaN	NaN	-3.521452
11	-0.238640	NaN	NaN	-3.302705
12	1.459323	NaN	NaN	-3.024274
13	-0.808426	NaN	NaN	-2.991200
14	-0.022121	NaN	NaN	-2.808623
15	0.034858	NaN	NaN	-2.773638
16	-0.193057	NaN	NaN	-2.757956
17	-0.626095	NaN	NaN	-2.570733
18	-0.500742	NaN	NaN	-2.514351
19	0.684414	NaN	NaN	-2.497775
20	1.288388	NaN	NaN	-2.477820
21	NaN	-0.742460	0.618333	-2.458530
22	-0.432367	NaN	NaN	-2.209709
23	-0.489346	NaN	NaN	-2.087824
24	-1.161694	-0.146058	-0.400098	-2.077006

```
[469]: res[res.Pos == "GK"]
```

```
[469]:
```

	Player Pos	Gls	Standard FK	Total Cmp%	Medium Cmp%	\
1	B. Ampofo GK	-0.118982	0.540325	0.321024	1.613995	
2	A. Omar GK	-0.903910	-0.551745	0.913974	1.610083	
3	M. Asiiimwe GK	-0.602015	-0.066381	0.185933	1.606172	
21	W. Nasiru GK	NaN	NaN	NaN	NaN	
24	F. Akumu GK	-0.058603	-1.401133	0.349130	1.486869	
44	W. Nasiru GK	-0.058603	-0.794427	-0.501309	1.422329	
54	F. Ithungu GK	-0.843531	-0.430404	-2.395303	1.041930	
57	Y. Acola GK	-1.024668	0.054960	0.371796	1.467312	
64	Y. Draru GK	NaN	NaN	NaN	NaN	
76	U. Nyeko GK	0.001776	-0.066381	-0.224780	1.388103	
86	Y. Draru GK	-0.722773	-0.915768	-1.095165	1.598348	
111	Y. Acola GK	NaN	NaN	NaN	NaN	
124	G. Leitner GK	-0.662394	0.783007	-0.270113	1.455577	
161	N. Irawan GK	0.001776	-0.551745	-1.229349	1.400815	
164	S. Marchetti GK	-0.420878	-1.279792	-2.618339	1.606172	
169	X. Tumushabe GK	-0.058603	-1.037109	-3.063504	1.611061	
171	Z. Nyamahunge GK	-0.300119	-1.401133	-0.559334	1.299115	

	Blocks Pass	Performance GA	Penalty Kicks PKm	Value
1	-1.150298	-1.147161	-0.909313	-8.482538
2	-1.252860	-1.573163	-0.909313	-8.252954
3	-1.230068	-1.253662	-0.739575	-7.419922
21	NaN	-0.742460	0.618333	-2.458530
24	-1.161694	-0.146058	-0.400098	-2.077006
44	-1.321234	NaN	NaN	-1.430853
54	-1.173090	-0.593359	0.618333	-1.135393
57	-1.309838	NaN	NaN	-1.060673
64	NaN	-0.465559	1.297287	-0.692210
76	-1.309838	0.514245	-0.060621	-0.240823
86	-1.116111	NaN	NaN	0.015992
111	NaN	-0.018257	1.127548	0.978324
124	-1.184485	0.940246	-0.400098	1.517425
161	-1.264255	0.897646	1.467025	3.642832
164	-1.116111	1.046747	-1.079052	3.746545
169	-1.150298	0.940246	0.957810	4.542296
171	-1.309838	1.600549	-1.588267	4.737982

```
[361]: display(res.loc[res.Pos.isin(["DF", "DFMF", "MFDF"])] .head(10))
display(res.loc[res.Pos.isin(["MF", "MFFW", "FWMF", "MFDF", "DFMF"])] .head(10))
display(res.loc[res.Pos.isin(["FW", "MFFW", "FWMF"])] .head(10))
display(res.loc[res.Pos.isin(["GK"])] .head(10))
```

	Player Pos	Gls	Standard FK	Total Cmp%	Medium Cmp%	\
18	T. Larsson DF	-0.118982	0.661666	1.418978	1.180790	
25	E. Bahrami DF	-0.843531	0.297643	1.476097	1.606172	
28	T. Nouri DF	0.303671	0.054960	1.224955	1.133852	
30	W. Guo DF	0.122534	1.147030	0.452488	0.480623	
34	C. Baluka DF	0.484808	0.297643	0.777976	0.636107	
36	I. Takeuchi DF	-0.662394	0.418984	-0.282806	1.604216	
37	H. Azizi DF	0.665946	-0.915768	0.767096	0.761277	
42	B. Lai DF	-0.602015	0.904348	1.069011	0.694780	
43	C. Tukamushaba DF	-0.179361	-0.187722	1.220422	1.047798	
45	R. Taketa DF	-0.360499	0.418984	1.049065	0.958810	

	Blocks Pass	Performance GA	Penalty Kicks PKm	Value
18	-0.500742	NaN	NaN	-2.514351
25	-1.218673	NaN	NaN	-2.062101
28	0.034858	NaN	NaN	-2.011924
30	-0.010725	NaN	NaN	-1.905112
34	0.183002	NaN	NaN	-1.774382
36	-1.138902	NaN	NaN	-1.695318
37	-0.546325	NaN	NaN	-1.686132
42	-0.637490	NaN	NaN	-1.517731
43	-0.534929	NaN	NaN	-1.444860
45	-0.352597	NaN	NaN	-1.415379

	Player	Pos	Gls	Standard FK	Total Cmp%	Medium Cmp%	\
4	J. Namirembe	MF	2.115043	2.117759	-0.915648	-0.898199	
6	P. Rabiw	MFFW	0.847083	4.059216	-0.465949	-0.853216	
8	D. Lehner	MFFW	2.477317	0.661666	-0.633680	-0.877663	
9	G. Jankowski	MF	2.537697	-0.309063	0.062628	-0.191186	
11	P. Villa	MFFW	1.632011	-0.309063	0.884961	0.451286	
12	H. Tourgeman	MFFW	0.967841	1.753736	1.068104	0.790613	
17	L. Mandala	MFFW	1.390494	-0.187722	-0.179448	0.018082	
19	O. Wanjala	MF	1.148978	0.054960	1.074451	0.886446	
20	B. Ayuba	MF	2.839592	0.418984	-0.094222	-1.419413	
22	B. Katooko	MF	-0.541636	4.180558	-0.372564	-1.259039	

	Blocks Pass	Performance GA	Penalty Kicks	PKm	Value
4	-0.375389	NaN		NaN	-4.578821
6	0.194398	NaN		NaN	-4.083036
8	-0.250036	NaN		NaN	-3.796656
9	0.114628	NaN		NaN	-3.533854
11	-0.238640	NaN		NaN	-3.302705
12	1.459323	NaN		NaN	-3.024274
17	-0.626095	NaN		NaN	-2.570733
19	0.684414	NaN		NaN	-2.497775
20	1.288388	NaN		NaN	-2.477820
22	-0.432367	NaN		NaN	-2.209709

	Player	Pos	Gls	Standard FK	Total Cmp%	Medium Cmp%	\
0	K. Kazlo?	FW	4.832101	3.937875	-0.578374	-0.882552	
5	H. Mubaiwa	FW	2.598076	0.054960	-0.253793	-0.867884	
6	P. Rabiw	MFFW	0.847083	4.059216	-0.465949	-0.853216	
7	U. Shoko	FW	2.718834	-0.551745	-0.018970	-0.073839	
8	D. Lehner	MFFW	2.477317	0.661666	-0.633680	-0.877663	
10	H. Makumbi	FW	2.175422	-0.673086	-0.062490	-0.201943	
11	P. Villa	MFFW	1.632011	-0.309063	0.884961	0.451286	
12	H. Tourgeman	MFFW	0.967841	1.753736	1.068104	0.790613	
13	L. Tarigan	FW	3.020729	-1.279792	-0.339925	-1.393010	
14	A. Kyarikunda	FW	0.967841	1.025689	0.388116	0.118805	

	Blocks Pass	Performance GA	Penalty Kicks	PKm	Value
0	-0.945175	NaN		NaN	-11.336899
5	-1.013549	NaN		NaN	-4.269473
6	0.194398	NaN		NaN	-4.083036
7	-0.136078	NaN		NaN	-3.895665
8	-0.250036	NaN		NaN	-3.796656
10	-0.956571	NaN		NaN	-3.521452
11	-0.238640	NaN		NaN	-3.302705
12	1.459323	NaN		NaN	-3.024274
13	-0.808426	NaN		NaN	-2.991200
14	-0.022121	NaN		NaN	-2.808623

Player	Pos	Gls	Standard FK	Total Cmp%	Medium Cmp%	\
--------	-----	-----	-------------	------------	-------------	---

1	B. Ampofo	GK	-0.118982	0.540325	0.321024	1.613995
2	A. Omar	GK	-0.903910	-0.551745	0.913974	1.610083
3	M. Asimwe	GK	-0.602015	-0.066381	0.185933	1.606172
21	W. Nasiru	GK	NaN	NaN	NaN	NaN
24	F. Akumu	GK	-0.058603	-1.401133	0.349130	1.486869
44	W. Nasiru	GK	-0.058603	-0.794427	-0.501309	1.422329
54	F. Ithungu	GK	-0.843531	-0.430404	-2.395303	1.041930
57	Y. Acola	GK	-1.024668	0.054960	0.371796	1.467312
64	Y. Draru	GK	NaN	NaN	NaN	NaN
76	U. Nyeko	GK	0.001776	-0.066381	-0.224780	1.388103

	Blocks Pass	Performance GA	Penalty Kicks PKm	Value
1	-1.150298	-1.147161	-0.909313	-8.482538
2	-1.252860	-1.573163	-0.909313	-8.252954
3	-1.230068	-1.253662	-0.739575	-7.419922
21	NaN	-0.742460	0.618333	-2.458530
24	-1.161694	-0.146058	-0.400098	-2.077006
44	-1.321234	NaN	NaN	-1.430853
54	-1.173090	-0.593359	0.618333	-1.135393
57	-1.309838	NaN	NaN	-1.060673
64	NaN	-0.465559	1.297287	-0.692210
76	-1.309838	0.514245	-0.060621	-0.240823

```
[508]: # selected_indices = [0,1,4,5,9,11,16,18,22,24,25,28,29,31,33,34,39,42,46,70]
selected_indices = [0,1,4,5,9,11,16,18,22,3,25,28,29,31,33,34,39,42,46,70]
# selected_indices = [0,1,4,5,9,11,16,18,24,25,26,28,29,31,34,39,42,46,52]
players = res.iloc[selected_indices,0]
players
```

```
[508]: 0      K. Kazlo?
1      B. Ampofo
4      J. Namirembe
5      H. Mubaiwa
9      G. Jankowski
11     P. Villa
16     N. Yamashita
18     T. Larsson
22     B. Katooko
3      M. Asimwe
25     E. Bahrami
28     T. Nouri
29     W. Dehghani
31     E. Nakanjako
33     K. Al-Zahrani
34     C. Baluka
39     H. Amade
42     B. Lai
```

46            H. Zare  
70            V. Sultan  
Name: Player, dtype: object

Rank determined based on Lasso Regression. Error term in Lasso assumed to be normally distributed with mean zero and constant variance (independent of predictors).

Let  $r$  denote the rank and  $\hat{r}$  denote the predicted rank. Correcting for continuous vs discrete, the probability of ranking first is

$$\Pr[r \leq 1.5] = \Pr\left[\frac{r - \hat{r}}{\sigma_r} \leq \frac{1.5 - \hat{r}}{\sigma_r}\right] = \Phi\left(\frac{1.5 - \hat{r}}{\sigma_r}\right)$$

where  $\Phi$  is the CDF of a standard normal random variable.

The variance  $\sigma^2$  is estimated from the data on  $r - \hat{r}$ .

The probability of top 10 is calculated similarly.

$$\Pr[r \leq 10.5]$$

```
[509]: # Predicting rank
value = [res.iloc[selected_indices].mean()[stat] * coef for stat, coef in_lasso_coefs.values]
rank = sum(value) + lasso.intercept_
print(rank)
```

4.1890246660050945

```
[510]: # Pr(rank <= 1.5) = Pr(Z <= (1.5-y_pred)/std)
z_rank_1 = (1.5 - rank) / np.std(y-y_pred)
z_rank_1
```

[510]: -1.0550332569310008

```
[511]: # Probability of rank 1 in 1 year.
norm.cdf(z_rank_1)
```

[511]: 0.14570505745874235

```
[512]: z_rank_10 = (10.5 - rank) / np.std(y-y_pred)
z_rank_10
```

[512]: 2.476098098024379

```
[513]: # Probability of top 10 in one year
norm.cdf(z_rank_10)
```

[513]: 0.99335864616137

```
[514]: # Probabilty of first in 10 years
1-(1-norm.cdf(z_rank_1))*10
```

```
[514]: 0.7929485132643551
```

```
[496]: res
```

```
[496]:
```

	Player	Pos	Gls	Standard FK	Total Cmp%	Medium Cmp%	\
0	K. Kazlo?	FW	4.832101	3.937875	-0.578374	-0.882552	
1	B. Ampofo	GK	-0.118982	0.540325	0.321024	1.613995	
2	A. Omar	GK	-0.903910	-0.551745	0.913974	1.610083	
3	M. Asiiimwe	GK	-0.602015	-0.066381	0.185933	1.606172	
4	J. Namirembe	MF	2.115043	2.117759	-0.915648	-0.898199	
..	...	...	...	...	...	...	
168	O. Balog	DFMF	-0.843531	-1.522474	-1.519478	-1.876086	
169	X. Tumushabe	GK	-0.058603	-1.037109	-3.063504	1.611061	
170	L. Jung	FWMF	-0.903910	-1.158451	-1.593823	-2.054062	
171	Z. Nyamahunge	GK	-0.300119	-1.401133	-0.559334	1.299115	
172	F. Andreassen	FWMF	-0.541636	-1.037109	-0.196674	-1.651172	

	Blocks Pass	Performance GA	Penalty Kicks	PKm	Value
0	-0.945175	NaN		NaN	-11.336899
1	-1.150298	-1.147161	-0.909313	-8.482538	
2	-1.252860	-1.573163	-0.909313	-8.252954	
3	-1.230068	-1.253662	-0.739575	-7.419922	
4	-0.375389	NaN		NaN	-4.578821
..	...	...	...	...	...
168	-0.637490	NaN		NaN	4.505589
169	-1.150298	0.940246	0.957810	4.542296	
170	-0.329806	NaN		NaN	4.726438
171	-1.309838	1.600549	-1.588267	4.737982	
172	1.710029	NaN		NaN	4.968167

```
[173 rows x 10 columns]
```

```
[497]: res.iloc[selected_indices]
```

```
[497]:
```

	Player	Pos	Gls	Standard FK	Total Cmp%	Medium Cmp%	\
0	K. Kazlo?	FW	4.832101	3.937875	-0.578374	-0.882552	
1	B. Ampofo	GK	-0.118982	0.540325	0.321024	1.613995	
4	J. Namirembe	MF	2.115043	2.117759	-0.915648	-0.898199	
5	H. Mubaiwa	FW	2.598076	0.054960	-0.253793	-0.867884	
9	G. Jankowski	MF	2.537697	-0.309063	0.062628	-0.191186	
11	P. Villa	MFFW	1.632011	-0.309063	0.884961	0.451286	
16	N. Yamashita	FW	3.322624	-0.187722	-0.846743	-2.499978	
18	T. Larsson	DF	-0.118982	0.661666	1.418978	1.180790	
22	B. Katooko	MF	-0.541636	4.180558	-0.372564	-1.259039	

3	M. Asiiimwe	GK	-0.602015	-0.066381	0.185933	1.606172
25	E. Bahrami	DF	-0.843531	0.297643	1.476097	1.606172
28	T. Nouri	DF	0.303671	0.054960	1.224955	1.133852
29	W. Dehghani	MF	-0.964289	0.176301	2.079927	1.608127
31	E. Nakanjako	MFFW	0.243292	0.418984	0.964746	0.778879
33	K. Al-Zahrani	FW	1.028220	0.540325	-0.260140	-0.580385
34	C. Baluka	DF	0.484808	0.297643	0.777976	0.636107
39	H. Amade	MFFW	0.907462	0.054960	0.340063	0.367188
42	B. Lai	DF	-0.602015	0.904348	1.069011	0.694780
46	H. Zare	DF	-0.783152	0.661666	1.490604	1.333341
70	V. Sultan	DFMF	-0.783152	0.661666	0.913974	0.954898

	Blocks Pass	Performance GA	Penalty Kicks PKm	Value
0	-0.945175	NaN	NaN	-11.336899
1	-1.150298	-1.147161	-0.909313	-8.482538
4	-0.375389	NaN	NaN	-4.578821
5	-1.013549	NaN	NaN	-4.269473
9	0.114628	NaN	NaN	-3.533854
11	-0.238640	NaN	NaN	-3.302705
16	-0.193057	NaN	NaN	-2.757956
18	-0.500742	NaN	NaN	-2.514351
22	-0.432367	NaN	NaN	-2.209709
3	-1.230068	-1.253662	-0.739575	-7.419922
25	-1.218673	NaN	NaN	-2.062101
28	0.034858	NaN	NaN	-2.011924
29	-1.207277	NaN	NaN	-1.984217
31	-0.022121	NaN	NaN	-1.890153
33	-0.329806	NaN	NaN	-1.840628
34	0.183002	NaN	NaN	-1.774382
39	0.342542	NaN	NaN	-1.663179
42	-0.637490	NaN	NaN	-1.517731
46	-0.272827	NaN	NaN	-1.405101
70	0.103232	NaN	NaN	-0.494604

```
[498]: tournament_dfs_2020 = []
for df in ["tournament_shooting", "tournament_passing", "tournament_defense",
           "tournament_goalkeeping"]:
    tournament_dfs_2020.append(dfs[df][dfs[df].Year==2020])

merged_df_2020 = reduce(lambda left, right: pd.merge(left, right,
                                                    how='outer',
                                                    on=["Player", "Nation", "Age", "Pos", "Born", "League", "Year"]),
                        tournament_dfs_2020)
merged_df_2020
```



[498]:

	Player	Nation	Pos	Age	Born	90s_x	Gls	Standard Sh	\
0	B. Schneider	Bernepamar	DF	26	1995	3.02	-0.06	-0.02	
1	G. Adilovi?	Bernepamar	DF	19	2002	2.92	-0.09	0.08	
2	J. Mizrahi	Bernepamar	DF	22	1999	1.95	0.04	0.55	
3	L. Ayugi	Bernepamar	DF	29	1992	1.96	0.02	0.02	
4	N. Salleh	Bernepamar	DF	24	1997	2.26	-0.07	0.37	
...	...	...	...	...	...	...	...	...	
1522	P. Uddin	Xikong	MF	30	1991	4.89	0.00	1.93	
1523	S. Yamamoto	Xikong	MF	36	1985	2.70	-0.02	0.45	
1524	X. Aliru	Xikong	MF	22	1999	0.11	0.01	10.00	
1525	X. Arab	Xikong	MF	34	1987	0.51	0.09	-0.01	
1526	U. Taaka	Xikong	MFFW	21	2000	5.00	0.11	1.49	

	Standard SoT	Standard SoT%	...	W	D	L	Performance CS	\
0	0.02	NaN	...	NaN	NaN	NaN	NaN	
1	0.06	NaN	...	NaN	NaN	NaN	NaN	
2	0.02	-0.04	...	NaN	NaN	NaN	NaN	
3	-0.05	NaN	...	NaN	NaN	NaN	NaN	
4	0.48	100.07	...	NaN	NaN	NaN	NaN	
...	...	...	...	...	...	...	...	
1522	0.76	44.39	...	NaN	NaN	NaN	NaN	
1523	0.03	0.04	...	NaN	NaN	NaN	NaN	
1524	-0.01	0.03	...	NaN	NaN	NaN	NaN	
1525	-0.07	NaN	...	NaN	NaN	NaN	NaN	
1526	0.51	42.95	...	NaN	NaN	NaN	NaN	

	Performance CS%	Performance PKatt_y	Penalty Kicks PKA	\
0	NaN	NaN	NaN	
1	NaN	NaN	NaN	
2	NaN	NaN	NaN	
3	NaN	NaN	NaN	
4	NaN	NaN	NaN	
...	...	...	...	
1522	NaN	NaN	NaN	
1523	NaN	NaN	NaN	
1524	NaN	NaN	NaN	
1525	NaN	NaN	NaN	
1526	NaN	NaN	NaN	

	Penalty Kicks PKsv	Penalty Kicks PKm	Penalty Kicks Save%
0	NaN	NaN	NaN
1	NaN	NaN	NaN
2	NaN	NaN	NaN
3	NaN	NaN	NaN
4	NaN	NaN	NaN
...	...	...	...
1522	NaN	NaN	NaN

1523	NaN	NaN	NaN
1524	NaN	NaN	NaN
1525	NaN	NaN	NaN
1526	NaN	NaN	NaN

[1527 rows x 91 columns]

```
[499]: df2 = merged_df_2020.groupby("Nation").mean()
df2["Number Players"] = merged_df_2020.groupby("Nation")["Player"].count()
df2 = df2.dropna(how="all",axis=1)
df2["TOP10"] = df2.index.isin(dfs["tournament_results"]["2021"][:10].dropna()).
↳astype("int")
# df2 = df.drop("Dosqaly") # No GK
df2 = df2.drop(["Year","Penalty Kicks Save%"], axis=1)
df2
```

```
[499]:
```

	Age	Born	90s_x	Gls \
Nation				
Bernepamar	27.192308	1993.692308	2.105385	0.066923
Bernoullist	27.476190	1993.523810	2.061429	0.002857
Byasier Pujan	27.833333	1992.958333	2.776250	0.067500
Cabballi	25.535714	1995.357143	2.386071	0.024286
Cabral Retrea	28.612903	1992.258065	1.720323	0.241935
Cuandbo	29.533333	1991.400000	2.185000	0.133667
Dastatesne	27.384615	1993.423077	2.457692	0.228077
Deshslands Landdenhai	28.250000	1992.625000	2.044062	0.015625
Djipines	25.750000	1995.071429	2.339643	0.094643
Dosqaly	26.733333	1994.066667	2.665667	0.179000
Eastern Niasland	28.444444	1992.481481	2.425556	0.085556
Eastern Sleboube	29.172414	1991.724138	2.258621	0.103103
Esia	26.785714	1994.214286	2.362143	0.121786
Galamily	28.651163	1992.302326	1.545814	0.139767
Giumle Lizeibon	28.400000	1992.520000	2.628800	0.633200
Greri Landmoslands	28.148148	1992.555556	2.438519	0.266667
Highhlaands	28.444444	1992.481481	2.424444	0.053704
Humberstonia	29.037037	1991.851852	2.460741	0.091111
Ili Siaco	28.120000	1992.800000	2.636400	0.156000
Ingre	26.827586	1994.000000	2.285862	0.048621
Isle of Jabber	28.035714	1992.857143	2.341071	0.056071
Isle of Lababwe	28.807692	1992.153846	2.555385	0.165385
Iyan	26.346154	1994.538462	2.540385	0.097308
Kani	28.103448	1992.827586	2.244138	0.008276
Kesternsri	26.241379	1994.620690	2.271724	0.048276
Klausterton	29.130435	1991.782609	2.805217	-0.014783
Ledian	27.939394	1992.818182	1.990606	0.070000
Leoneku Guidisia	28.038462	1992.884615	2.498077	0.099231
Liacra	26.047619	1994.809524	3.128571	0.052381

Loco Phirema	26.851852	1994.111111	2.452222	0.087407
Mandlestan	26.454545	1994.454545	3.014545	0.522727
Manlisgamncent	27.518519	1993.259259	2.422593	0.099630
Mico	27.640000	1993.320000	2.619200	0.084400
Moaithe	26.774194	1994.064516	2.136774	0.078065
Naguayli	27.807692	1993.076923	2.543846	0.036154
New Uwi	27.800000	1993.040000	2.655600	0.150000
Nganion	26.277778	1994.638889	2.452222	0.128889
Ngoque Blicri	28.392857	1992.500000	2.346786	0.059643
Nkasland Cronestan	27.583333	1993.416667	2.744167	0.089583
Pahon	28.640000	1992.200000	2.639600	0.146400
People's Land of Maneau	25.206897	1995.655172	2.261034	0.104483
Prometricia	26.909091	1994.000000	1.996364	0.059091
Quewenia	27.041667	1993.791667	2.261667	0.160000
Redohrainbri	26.692308	1994.153846	1.402051	0.103333
Rosvi	27.214286	1993.607143	2.355714	0.142143
Saintu	28.080000	1992.720000	2.650000	0.079600
Sobianitedrucy	27.593750	1993.312500	2.720937	0.118125
Southern Ristan	27.785714	1993.107143	3.152143	0.249286
Southslands	27.344828	1993.517241	2.104483	0.004483
Tiagascar Westlands	27.709677	1993.225806	2.113871	0.068710
Tiliqoiuy	26.275862	1994.620690	2.251379	0.009655
Unicorporated Tiagascar	27.441176	1993.441176	1.904118	0.053824
Varijitri Isles	26.533333	1994.300000	2.176000	0.049333
West Iyan	28.454545	1992.409091	1.986818	0.043182
Xikong	29.000000	1991.960000	2.601600	0.066800

	Standard Sh	Standard SoT	Standard SoT% \
Nation			
Bernepamar	0.598462	0.219615	34.594667
Bernoullist	0.638095	0.168095	19.697500
Byasier Pujan	1.724583	0.437917	28.817368
Cabballi	0.738571	0.170357	29.183333
Cabral Retrea	1.141935	0.489677	35.546190
Cuandbo	0.469000	0.318333	51.106667
Dastatesne	1.143462	0.438846	26.537500
Deshslands Landdenhai	0.643125	0.140313	29.656429
Djipines	1.107143	0.353214	32.888824
Dosqaly	0.827333	0.319667	39.382353
Eastern Niasland	1.464815	0.385556	27.241000
Eastern Sleboube	1.427931	0.410000	33.801579
Esia	0.974643	0.452143	54.244444
Galamily	1.150000	0.292791	24.454583
Giumle Lizeibon	2.166400	0.887600	40.811053
Greri Landmoslands	1.637037	0.554074	32.475000
Highhlaands	0.500741	0.194444	39.760000
Humberstonia	0.835926	0.252222	33.837368

Ili Siaco	1.075600	0.554800	42.334706
Ingre	1.313793	0.412759	31.356842
Isle of Jabber	0.834286	0.197500	29.567059
Isle of Lababwe	1.148077	0.326154	33.645000
Iyan	1.181538	0.325000	30.904737
Kani	1.114138	0.137931	30.942857
Kesternsri	0.586207	0.150000	24.727059
Klausterton	0.629130	0.077391	23.210769
Ledian	0.697576	0.273333	32.863889
Leoneku Guidisia	0.978846	0.245769	22.933684
Liacra	1.189524	0.372857	28.272857
Loco Phirema	1.020741	0.445185	42.841500
Mandlestan	1.119545	0.707273	36.299286
Manlisgamncent	0.669630	0.226296	34.691429
Mico	0.601200	0.174000	24.620909
Moaithe	1.491935	0.647097	36.126087
Naguayli	0.942692	0.241154	18.050000
New Uwi	0.922000	0.416400	41.729286
Nganion	1.300833	0.486111	31.532273
Ngoque Blicri	1.095000	0.419643	18.773158
Nkasland Cronestan	1.076667	0.307917	38.648667
Pahon	0.941200	0.365200	34.053125
People's Land of Maneau	1.256207	0.415172	33.510000
Prometricia	1.070455	0.230000	30.127500
Quewenia	1.113333	0.450833	39.417647
Redohrainbri	0.737436	0.354359	43.187222
Rosvi	1.471071	0.536071	27.153333
Saintu	0.771200	0.282800	39.178125
Sobianitedrucy	1.684375	0.259063	23.535000
Southern Ristan	0.715714	0.382500	47.067647
Southslands	1.357241	0.458621	25.814000
Tiagascar Westlands	1.659677	0.438387	25.498000
Tiliqoiuy	1.148966	0.428966	33.372222
Unicorporated Tiagascar	0.520882	0.156176	19.038750
Varijitri Isles	1.134000	0.417000	30.191053
West Iyan	0.585000	0.136364	24.988889
Xikong	1.406000	0.191600	24.554706

	Standard Sh/90	Standard SoT/90	Standard G/Sh	...	\
Nation				...	
Bernepamar	0.610769	0.183846	0.086667	...	
Bernoullist	0.584762	0.132381	0.013333	...	
Byasier Pujan	1.947083	0.435833	0.047895	...	
Cabballi	0.747500	0.181429	-0.014167	...	
Cabral Retrea	1.149032	0.487419	0.107619	...	
Cuandbo	0.455000	0.310000	0.139333	...	
Dastatesne	1.163462	0.447692	0.080625	...	

Deshslands Landddenhai	0.625625	0.161250	0.018571 ...
Djipines	1.085714	0.350357	0.138824 ...
Dosqaly	0.870667	0.315667	0.196471 ...
Eastern Niasland	1.591111	0.388148	0.055500 ...
Eastern Sleboube	1.396552	0.433448	0.137895 ...
Esia	0.996429	0.450357	0.153333 ...
Galamily	1.192558	0.296977	0.060417 ...
Giumle Lizeibon	2.003200	0.758800	0.148421 ...
Greri Landmoslands	1.653333	0.570741	0.079444 ...
Highhlaands	0.522222	0.199259	0.097500 ...
Humberstonia	0.826296	0.260000	0.130000 ...
Ili Siaco	1.083600	0.555600	0.109412 ...
Ingre	1.279655	0.402759	0.036316 ...
Isle of Jabber	0.806071	0.205000	0.138824 ...
Isle of Lababwe	1.187692	0.306538	0.161250 ...
Iyan	1.192692	0.327308	0.074211 ...
Kani	1.264828	0.151379	0.016667 ...
Kesternsri	0.594483	0.157241	0.049412 ...
Klausterton	0.615217	0.070000	0.015385 ...
Ledian	0.698182	0.249091	0.087778 ...
Leoneku Guidisia	0.974231	0.243077	0.090526 ...
Liacra	1.211905	0.393333	0.036429 ...
Loco Phirema	1.061481	0.447037	0.111500 ...
Mandlestan	1.110909	0.716818	0.196429 ...
Manlisgamncent	0.647407	0.228519	0.140714 ...
Mico	0.605200	0.180800	0.096364 ...
Moaithe	1.470968	0.627742	0.066522 ...
Naguayli	0.937308	0.226923	0.032941 ...
New Uwi	0.946000	0.399200	0.132143 ...
Nganion	1.319167	0.504444	0.132273 ...
Ngoque Blicri	1.102143	0.425000	0.035789 ...
Nkasland Cronestan	1.073333	0.302917	0.089333 ...
Pahon	0.917200	0.366400	0.088125 ...
People's Land of Maneau	1.304138	0.424138	0.060556 ...
Prometricia	1.085455	0.241364	0.066875 ...
Quewenia	1.122083	0.453750	0.095882 ...
Redohrainbri	0.715128	0.341282	0.102778 ...
Rosvi	1.483571	0.547143	0.056667 ...
Saintu	1.627200	1.156000	0.032500 ...
Sobianitedrucy	1.705313	0.270937	0.084545 ...
Southern Ristan	0.689286	0.383571	0.252353 ...
Southslands	1.343793	0.454483	-0.005000 ...
Tiagascar Westlands	1.823871	0.439355	0.028000 ...
Tiliqoiuy	1.192759	0.473448	0.019444 ...
Unicorporated Tiagascar	0.542353	0.163824	0.040625 ...
Varijitri Isles	1.134333	0.434333	0.026842 ...
West Iyan	0.585909	0.151364	0.064444 ...

Xikong 1.425200 0.216000 0.060588 ...

	D	L	Performance CS \
Nation			
Bernepamar	-4.000000e-02	0.660000	0.280000
Bernoullist	6.450000e-01	0.310000	0.625000
Byasier Pujan	2.500000e-01	0.135000	0.280000
Cabballi	1.300000e-01	0.675000	0.145000
Cabral Retrea	4.100000e-01	0.490000	0.040000
Cuandbo	-2.000000e-02	1.010000	-0.003333
Dastatesne	3.500000e-01	0.380000	0.090000
Deshslands Landdenhai	2.366667e-01	0.723333	0.130000
Djipines	1.400000e-01	0.046667	3.806667
Dosqaly	NaN	NaN	NaN
Eastern Niasland	8.750000e-01	0.110000	-0.025000
Eastern Sleboube	1.533333e-01	0.416667	0.016667
Esia	7.200000e-01	0.060000	0.043333
Galamily	2.891206e-18	0.423333	0.823333
Giumle Lizeibon	1.350000e-01	0.265000	0.255000
Greri Landmoslands	1.350000e-01	0.165000	0.475000
Highhlaands	2.000000e-01	0.756667	-0.026667
Humberstonia	1.750000e-01	0.300000	0.315000
Ili Siaco	3.050000e-01	0.055000	0.170000
Ingre	4.350000e-01	0.040000	0.745000
Isle of Jabber	1.900000e-01	0.660000	0.360000
Isle of Lababwe	3.500000e-01	0.180000	0.420000
Iyan	5.650000e-01	0.125000	0.420000
Kani	2.400000e-01	0.755000	0.070000
Kesternsri	4.000000e-01	0.540000	0.260000
Klausterton	2.900000e-01	0.680000	0.270000
Ledian	1.666667e-01	0.756667	0.083333
Leoneku Guidisia	1.000000e-02	0.280000	0.410000
Liacra	5.000000e-01	0.185000	0.580000
Loco Phirema	2.350000e-01	0.495000	0.380000
Mandlestan	5.100000e-01	0.200000	0.290000
Manlisgamncent	1.000000e-02	0.870000	0.020000
Mico	1.800000e-01	0.135000	0.305000
Moaithe	2.350000e-01	0.280000	0.445000
Naguayli	1.550000e-01	0.560000	0.215000
New Uwi	3.750000e-01	0.190000	0.675000
Nganion	2.750000e-01	0.180000	0.355000
Ngoque Blicri	7.500000e-02	0.520000	0.180000
Nkasland Cronestan	8.500000e-02	0.510000	0.305000
Pahon	2.100000e-01	0.240000	0.760000
People's Land of Maneau	1.000000e-01	0.310000	0.420000
Prometricia	6.600000e-01	0.215000	0.345000
Quewenia	7.000000e-01	0.330000	0.060000

Redohrainbri	3.266667e-01	0.100000	0.510000
Rosvi	4.250000e-01	0.325000	0.300000
Saintu	2.150000e-01	-0.065000	0.530000
Sobianitedrucy	2.800000e-01	0.030000	0.500000
Southern Ristan	3.666667e-02	0.250000	0.246667
Southslands	4.700000e-01	0.570000	0.310000
Tiagascar Westlands	6.450000e-01	0.260000	0.235000
Tiliqoiuy	5.000000e-01	0.370000	0.610000
Unicorporated Tiagascar	1.166667e-01	0.833333	0.090000
Varijitri Isles	3.500000e-01	0.635000	0.125000
West Iyan	3.500000e-01	0.045000	0.830000
Xikong	-4.000000e-02	0.800000	-0.010000

	Performance CS%	Performance PKatt_y \
Nation		
Bernepamar	24.935000	0.145000
Bernoullist	66.580000	0.000000
Byasier Pujan	25.000000	-0.070000
Cabballi	16.670000	0.145000
Cabral Retrea	0.090000	0.010000
Cuandbo	0.003333	0.576667
Dastatesne	16.700000	0.110000
Deshslands Landdenhai	11.086667	-0.046667
Djipines	75.075000	-0.016667
Dosqaly	NaN	NaN
Eastern Niasland	0.030000	0.210000
Eastern Sleboube	12.485000	0.233333
Esia	-0.036667	0.063333
Galamily	19.985000	0.283333
Giumle Lizeibon	25.010000	0.010000
Greri Landmoslands	49.975000	-0.080000
Highhlaands	-0.006667	0.053333
Humberstonia	33.365000	0.195000
Ili Siaco	20.015000	0.030000
Ingre	75.005000	-0.080000
Isle of Jabber	33.300000	0.230000
Isle of Lababwe	33.220000	0.470000
Iyan	37.500000	0.025000
Kani	0.015000	0.435000
Kesternsri	16.620000	-0.070000
Klausterton	33.240000	0.680000
Ledian	11.096667	0.186667
Leoneku Guidisia	49.950000	0.300000
Liacra	59.990000	0.220000
Loco Phirema	37.460000	0.250000
Mandlestan	33.300000	0.060000
Manlisgamncent	-0.060000	0.220000

Mico	49.920000	0.010000
Moaithe	50.045000	0.010000
Naguayli	19.980000	0.025000
New Uwi	66.625000	0.045000
Nganion	37.510000	-0.045000
Ngoque Blicri	19.970000	0.025000
Nkasland Cronestan	33.335000	0.560000
Pahon	83.360000	0.040000
People's Land of Maneau	49.920000	0.500000
Prometricia	33.300000	-0.085000
Quewenia	0.040000	0.470000
Redohrainbri	44.466667	0.166667
Rosvi	33.355000	0.140000
Saintu	50.045000	0.050000
Sobianitedrucy	50.070000	-0.010000
Southern Ristan	19.443333	0.250000
Southslands	33.260000	-0.100000
Tiagascar Westlands	25.035000	0.375000
Tiliqoiuy	66.640000	0.340000
Unicorporated Tiagascar	8.343333	-0.013333
Varijitri Isles	10.015000	0.115000
West Iyan	83.295000	-0.080000
Xikong	0.090000	0.100000

	Penalty Kicks PKA	Penalty Kicks PKsv \
Nation		
Bernepamar	0.210000	-0.005000
Bernoullist	-0.025000	0.065000
Byasier Pujan	0.000000	0.030000
Cabballi	0.040000	0.080000
Cabral Retrea	0.070000	-0.100000
Cuandbo	0.613333	0.013333
Dastatesne	0.170000	-0.020000
Deshslands Landdenhai	-0.006667	-0.006667
Djipines	0.043333	0.016667
Dosqaly	NaN	NaN
Eastern Niasland	0.305000	0.035000
Eastern Sleboube	0.263333	-0.016667
Esia	0.023333	-0.006667
Galamily	0.343333	-0.050000
Giumle Lizeibon	-0.070000	-0.045000
Greri Landmoslands	-0.055000	-0.065000
Highhlaands	-0.050000	-0.043333
Humberstonia	0.185000	0.085000
Ili Siaco	0.005000	0.035000
Ingre	-0.055000	0.000000
Isle of Jabber	0.070000	-0.100000



Isle of Lababwe	0.510000	0.080000
Iyan	0.000000	0.040000
Kani	0.385000	0.000000
Kesternsri	0.010000	0.020000
Klausterton	0.290000	0.290000
Ledian	0.113333	-0.003333
Leoneku Guidisia	0.190000	0.080000
Liacra	0.190000	-0.050000
Loco Phirema	0.195000	-0.010000
Mandlestan	-0.060000	0.050000
Manlisgamncent	0.090000	0.020000
Mico	-0.075000	-0.045000
Moaithe	0.025000	0.015000
Naguayli	0.175000	0.010000
New Uwi	0.010000	-0.010000
Nganion	0.060000	0.015000
Ngoque Blicri	-0.095000	-0.080000
Nkasland Cronestan	0.490000	0.045000
Pahon	-0.010000	0.060000
People's Land of Maneau	0.350000	0.060000
Prometricia	0.010000	0.070000
Quewenia	-0.050000	0.500000
Redohrainbri	0.120000	-0.010000
Rosvi	0.140000	-0.075000
Saintu	0.005000	-0.080000
Sobianitedrucy	0.090000	0.050000
Southern Ristan	0.250000	-0.040000
Southslands	0.040000	0.090000
Tiagascar Westlands	0.360000	0.025000
Tiliqoiuy	0.095000	0.015000
Unicorporated Tiagascar	-0.006667	0.036667
Varijitri Isles	0.075000	0.045000
West Iyan	0.005000	-0.020000
Xikong	0.050000	0.040000

	Penalty Kicks PKm	Number Players	TOP10
Nation			
Bernepamar	-0.050000	26	1
Bernoullist	-0.015000	21	0
Byasier Pujan	0.040000	24	0
Cabballi	0.255000	28	0
Cabral Retrea	0.050000	31	0
Cuandbo	-0.016667	30	0
Dastatesne	0.060000	26	0
Deshslands Landdenhai	-0.046667	32	0
Djipines	-0.016667	28	0
Dosqaly	NaN	30	1

Eastern Niasland	-0.060000	27	0
Eastern Sleboubé	0.073333	29	0
Esia	0.003333	28	0
Galamily	0.043333	43	1
Giumle Lizeibon	-0.030000	25	1
Greri Landmoslands	0.040000	27	0
Highhlaands	0.053333	27	0
Humberstonia	-0.015000	27	0
Ili Siaco	-0.030000	25	0
Ingre	0.010000	29	0
Isle of Jabber	-0.050000	28	0
Isle of Lababwe	0.030000	26	0
Iyan	0.035000	26	0
Kani	0.005000	29	0
Kesternsri	-0.040000	29	0
Klausterton	0.040000	23	0
Ledian	0.023333	33	0
Leoneku Guidisia	0.240000	26	0
Liacra	-0.035000	21	0
Loco Phirema	0.160000	27	0
Mandlestan	0.010000	22	0
Manlisgamncent	-0.080000	27	0
Mico	-0.025000	25	1
Moaithe	0.005000	31	0
Naguayli	-0.050000	26	0
New Uwi	0.005000	25	0
Nganion	-0.040000	36	1
Ngoque Blicri	0.015000	28	0
Nkasland Cronestan	0.020000	24	0
Pahon	-0.010000	25	0
People's Land of Maneau	0.220000	29	1
Prometricia	0.050000	22	0
Quewenia	0.050000	24	1
Redohrainbri	0.050000	39	0
Rosvi	0.075000	28	0
Saintu	0.060000	25	0
Sobianitedrucy	0.100000	32	1
Southern Ristan	0.036667	28	1
Southslands	-0.020000	29	0
Tiagascar Westlands	0.055000	31	0
Tiliqoiuy	0.230000	29	0
Unicorporated Tiagascar	0.020000	34	0
Varijitri Isles	-0.060000	30	0
West Iyan	0.035000	22	0
Xikong	0.250000	25	0

[55 rows x 33 columns]

```
[500]: lasso_coefs
```

```
[500]:
```

	Stat	Coef
0	Gls	-1.606882
1	Standard FK	-0.948223
2	Total Cmp%	-0.397181
3	Medium Cmp%	-0.896960
4	Blocks Pass	0.909493
5	Performance GA	4.089069
6	Penalty Kicks PKm	0.933866

```
[501]: test = df2[["Gls","Performance GA","Penalty Kicks PKm"]]
test = (test - test.mean()) / test.std()
test["Value"] = lasso.intercept_ + test["Gls"] * -1.606882 + test["Performance_
↪GA"] * 4.089069 + test["Penalty Kicks PKm"] * 0.933866
sorted_df = test.sort_values("Value")
```

```
[502]: for i,country in enumerate(dfs["tournament_results"]["2020"]):
sorted_df.loc[country,"Rank"] = i+1
```

```
[503]: sorted_df.dropna()
```

```
[503]:
```

	Gls	Performance GA	Penalty Kicks PKm \
Nation			
Nganion	0.156871	-1.267535	-0.901059
Sobianitedrucy	0.059907	-1.445587	0.837437
Greri Landmoslands	1.398018	-0.679189	0.092367
Southern Ristan	1.241445	-0.212124	0.050974
Byasier Pujan	-0.396140	-0.679189	0.092367
People's Land of Maneau	-0.062987	-0.655964	2.327576
Nkasland Cronestan	-0.197206	-0.083101	-0.155989
Mico	-0.243899	0.629108	-0.714792
Esia	0.092884	0.706522	-0.362953
Quewenia	0.437130	0.722005	0.216545
Bernepamar	-0.401337	1.085851	-1.025237
Unicorporated Tiagascar	-0.519342	1.155524	-0.155989
Manlisgamncent	-0.106706	2.518010	-1.397772
Xikong	-0.402445	1.620008	2.700111
Cuandbo	0.199911	3.008299	-0.611310

	Value	Rank
Nation		
Nganion	6.375593	2.0
Sobianitedrucy	7.426858	3.0
Greri Landmoslands	7.714733	5.0
Southern Ristan	9.837532	4.0
Byasier Pujan	10.597733	6.0

People's Land of Maneau	12.244748	8.0
Nkasland Cronestan	12.483583	10.0
Mico	14.949039	7.0
Esia	15.052990	9.0
Quewenia	15.104310	11.0
Bernepamar	16.779760	14.0
Unicorporated Tiagascar	18.066038	15.0
Manlisgamncent	21.814623	12.0
Xikong	22.444722	13.0
Cuandbo	24.061202	16.0

```
[504]: league_stats = leagues_merged_df.groupby("League").mean()
standard_league = (league_stats - league_stats.mean()) / league_stats.std()
```

```
[505]: standard_league[["Standard SoT/90", "Standard SoT%", "Medium Cmp%", "Pressures_
↳%", "Blocks Pass", "Performance GA90"]]
```

```
[505]:
```

	Standard SoT/90	Standard SoT%	Medium Cmp%	Pressures %	Blocks Pass \
League					
A	1.204254	1.731767	0.973231	-0.813712	0.318404
B	0.560981	-0.302532	0.335096	-0.367055	-1.358626
C	0.448441	0.298385	-1.426231	1.477211	1.414361
D	-0.630523	-1.277067	0.776450	-0.039744	0.205456
E	-1.623209	-0.435069	-1.066389	-1.124960	0.359109
RFL	0.040057	-0.015485	0.407843	0.868260	-0.938705

	Performance GA90
League	
A	-1.184309
B	0.163458
C	1.048021
D	-0.324239
E	-0.929215
RFL	1.226284

```
[ ]:
```