Import Libraries

In [1]:

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
```

Import Dataset

In [2]:

```
df_team = pd.read_csv('Teams.csv')
df_player = pd.read_csv('Players.csv')
```

Correlations (Teams.csv)

In [3]:

```
df_team.head(5)
```

Out[3]:

	team	ranking	games	wins	draws	losses	goalsFor	goalsAgainst	yellowCards	redCards
0	Brazil	1	5	3	1	1	9	4	7	2
1	Spain	2	6	5	0	1	7	2	3	0
2	Portugal	3	4	1	2	1	7	1	8	1
3	Netherlands	4	6	6	0	0	12	5	15	0
4	Italy	5	3	0	2	1	4	5	5	0

In [4]:

```
#correlation in team dataset
corr_team = df_team.corr()
corr_team
```

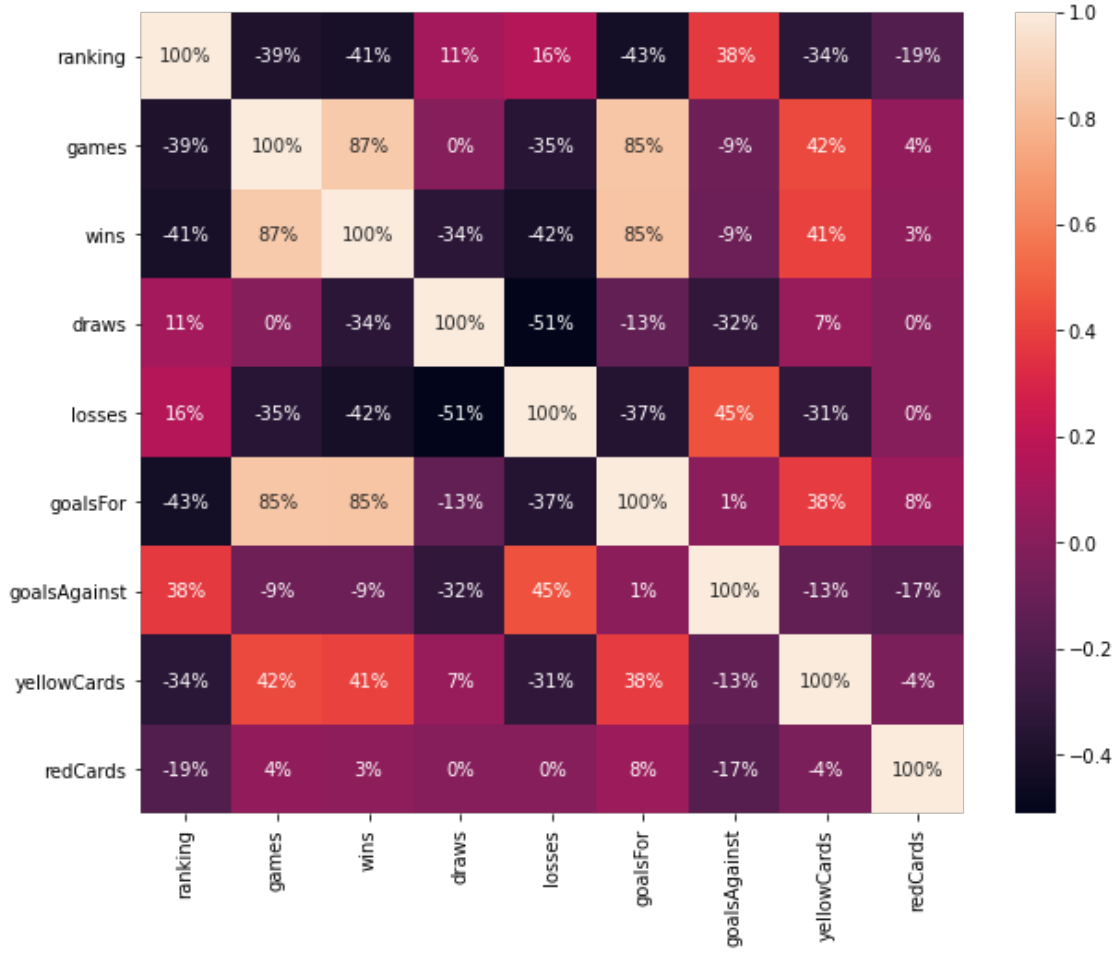
Out[4]:

	ranking	games	wins	draws	losses	goalsFor	goalsAgainst	yellowCards	redCards
ranking	1.000000	-0.394166	-0.413884	0.107413	0.159159	-0.430037	0.375879	-0.341489	-0.191772
games	-0.394166	1.000000	0.866114	0.000000	-0.347522	0.846791	-0.090421	0.421791	0.041959
wins	-0.413884	0.866114	1.000000	-0.337080	-0.418160	0.845161	-0.094121	0.405313	0.029488
draws	0.107413	0.000000	-0.337080	1.000000	-0.512316	-0.133786	-0.315388	0.066406	0.000000
losses	0.159159	-0.347522	-0.418160	-0.512316	1.000000	-0.367550	0.452844	-0.313332	0.000000
goalsFor	-0.430037	0.846791	0.845161	-0.133786	-0.367550	1.000000	0.014296	0.381244	0.075100
goalsAgainst	0.375879	-0.090421	-0.094121	-0.315388	0.452844	0.014296	1.000000	-0.130788	-0.172805
yellowCards	-0.341489	0.421791	0.405313	0.066406	-0.313332	0.381244	-0.130788	1.000000	-0.040665
redCards	-0.191772	0.041959	0.029488	0.000000	0.000000	0.075100	-0.172805	-0.040665	1.000000

Correlation Heatmap

In [5]:

```
plt.figure(figsize = (10,8))
ax = sns.heatmap(corr_team, annot = True, fmt = '.0%')
```



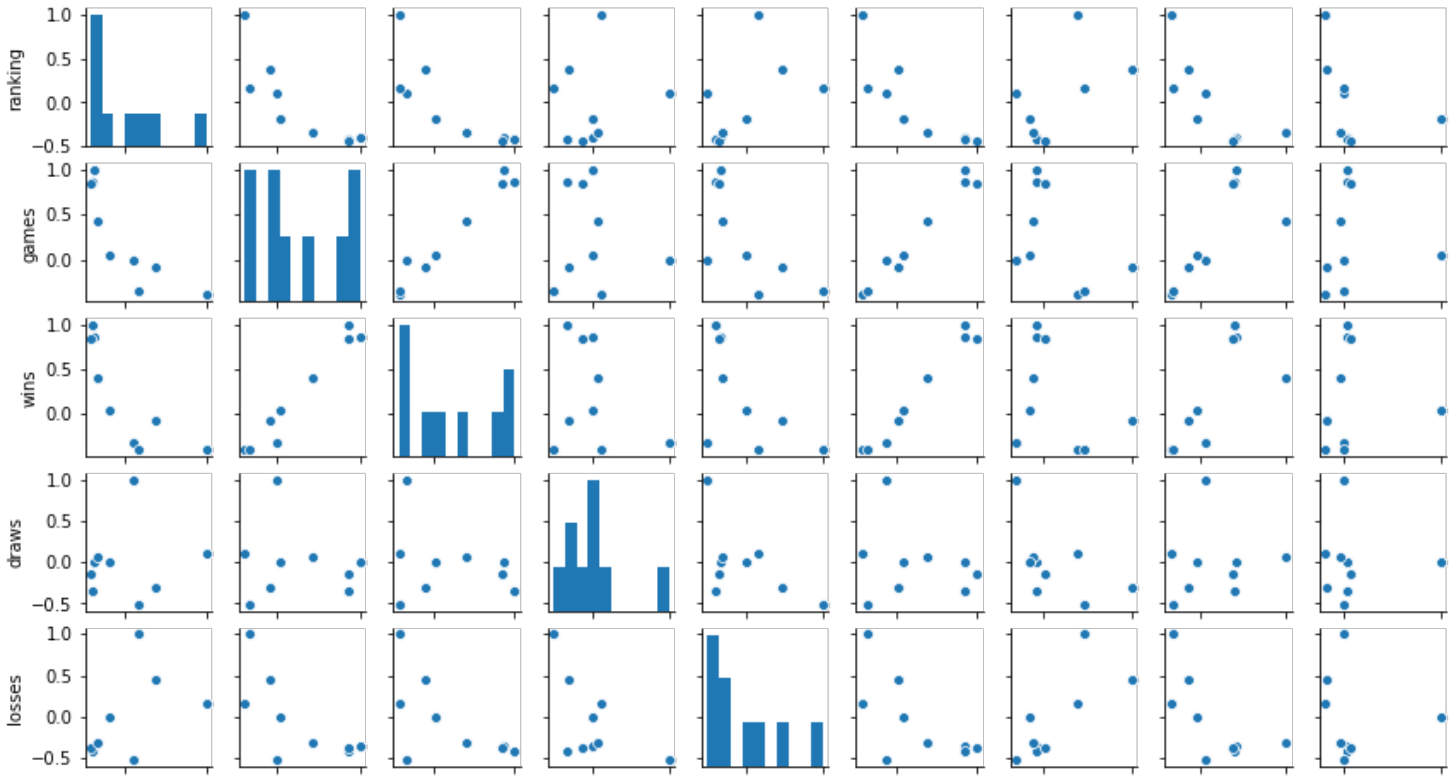
Seaborn Pairplot

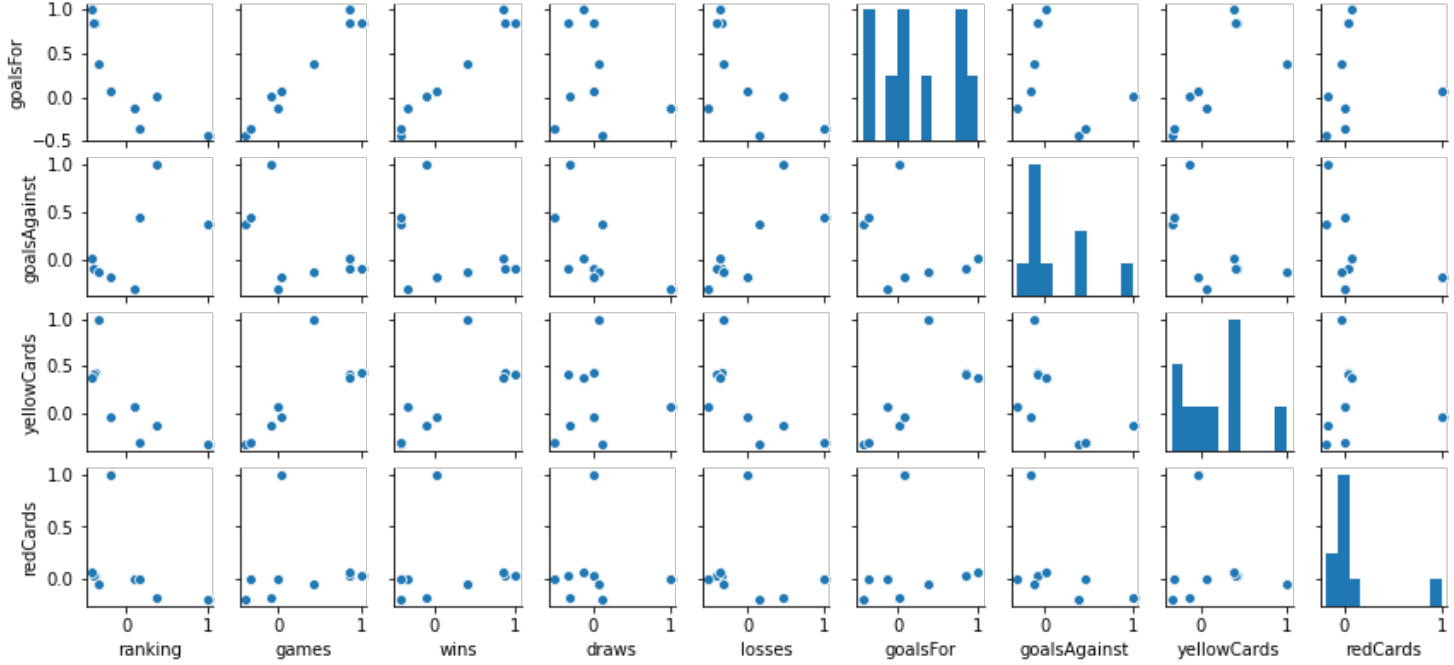
In [6]:

```
#compare pairplot and heatmap positive and negative correlation
sns.pairplot (corr_team, height = 1.3)
```

Out[6]:

```
<seaborn.axisgrid.PairGrid at 0x24c1162c640>
```





Correlations (Players.csv)

In [7]:

```
df_player.head(5)
```

Out[7]:

	surname	team	position	minutes	shots	passes	tackles	saves
0	Abdoun	Algeria	midfielder	16	0	6	0	0
1	Belhadj	Algeria	defender	270	1	146	8	0
2	Boudebouz	Algeria	midfielder	74	3	28	1	0
3	Bougherra	Algeria	defender	270	1	89	11	0
4	Chaouchi	Algeria	goalkeeper	90	0	17	0	2

In [15]:

```
#correlation in team dataset
corr_player = df_player.corr()
corr_player
```

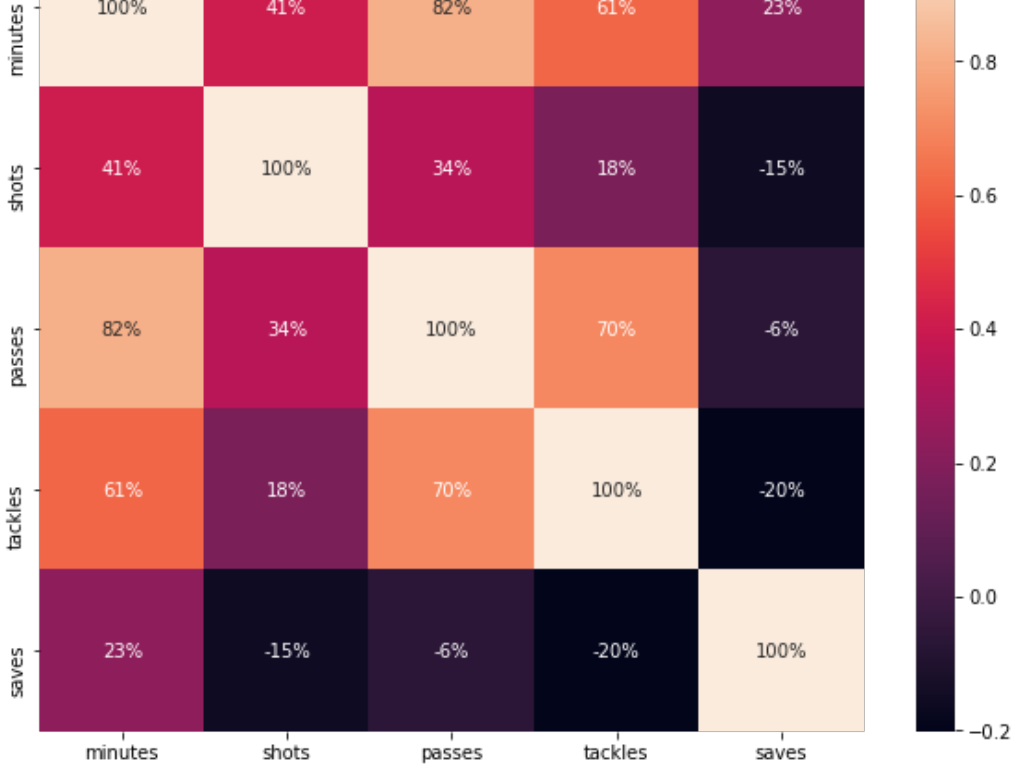
Out[15]:

	minutes	shots	passes	tackles	saves
minutes	1.000000	0.407923	0.815119	0.610673	0.228477
shots	0.407923	1.000000	0.343163	0.176283	-0.154958
passes	0.815119	0.343163	1.000000	0.702097	-0.062057
tackles	0.610673	0.176283	0.702097	1.000000	-0.201190
saves	0.228477	-0.154958	-0.062057	-0.201190	1.000000

Correlation Heatmap

In [9]:

```
plt.figure(figsize = (10,8))
ay = sns.heatmap(corr_player, annot = True, fmt = '.0%')
```



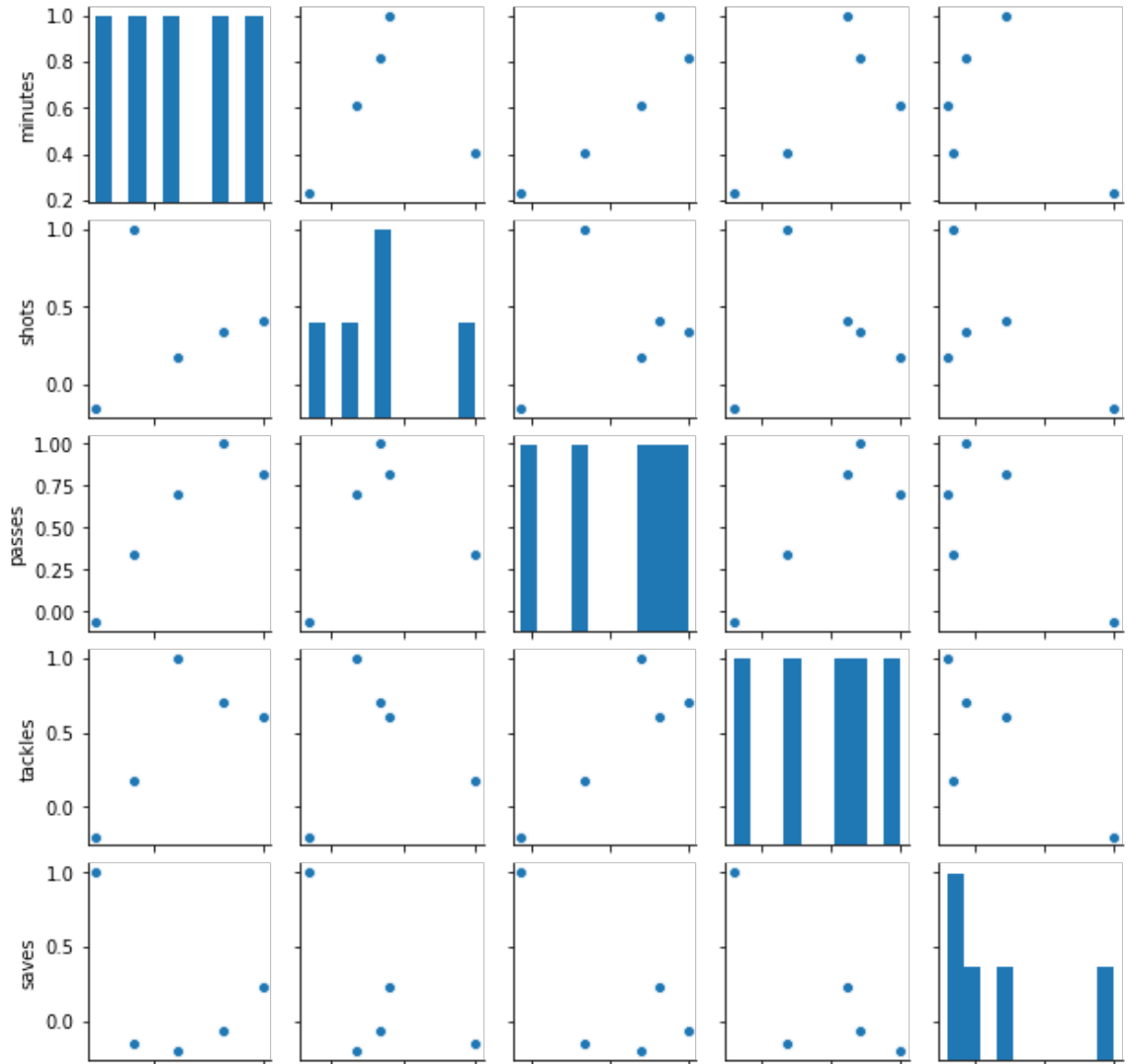
Seaborn Pairplot

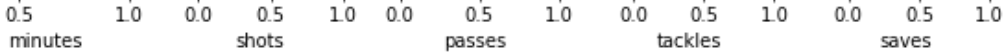
In [10]:

```
#compare pairplot and heatmap positive and negative correlation
sns.pairplot (corr_player, height = 1.7)
```

Out[10]:

```
<seaborn.axisgrid.PairGrid at 0x24c12ddbf40>
```




0.5
1.0
0.0
0.5
1.0
0.0
0.5
1.0
0.0
0.5
1.0
0.0
0.5
1.0
minutes
shots
passes
tackles
saves

In [ ]: