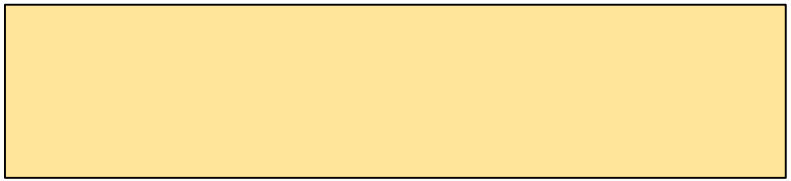


MINI PROJECT

Sai Durga Karthik Nandiraju

Soccer Dataset

I used Soccer Dataset for this mini project. The Soccer dataset contains 42 soccer player attributes like overall rating, vision, penalties strengths etc.

Motivation

In soccer games penalties play a crucial role in determining the winner. A team with penalties can loose players with good skills during match hours thus effecting the overall performance of the team. During a penalty the opposite team has a better chance to score a goal. Thus penalty is a problem sometimes for a soccer team. A defender in a soccer field is more likely to involve in penalties as he/she must prevent the other team from scoring. The insights of which player is more likely to associate with penalties based on his defensive work rate and which foot preferred player has a better overall rating would be useful to a coach or owner of the football team during player selection, training and strategy building.

Research Question(s)

1.Do the players with high defensive work rate have more penalties than the players with low defensive work rate?

2.Do the players with preferred right foot have a better overall rating than the preferred left foot players?

3.Do the players with low vision have more penalties than the players with high vision?

Findings

1. I found that on an average players with low defensive work rate are associated with more penalties than those with medium defensive work rate.

   Low defensive work rate average penalties: 64.76886

   Medium defensive work rate average penalties : 53.39736

2. I found that on an average both preferred right foot(prf) and preferred left foot(plf) players have same overall rating(Ort).

   prf Ort average: 68.6296 plf Ort average: 68.652

I attached a picture of my Jupyter notebook in the next slide.

Findings

```
avg_right= df.loc[df['preferred_foot'] == 1, 'overall_rating'].copy()
```

```
avg_right.mean()
```

```
68.62965056111327
```

```
avg_left= df.loc[df['preferred_foot'] == 0, 'overall_rating'].copy()
```

```
avg_left.mean()
```

```
68.65282154760015
```

Defensive_work_rate= "low" is labeled 0 and "medium" is labeled 1

```
avg_penalties_lowdefense= df.loc[df['defensive_work_rate'] == 0, 'penalties'].copy()
```

```
avg_penalties_lowdefense.mean()
```

```
64.76866319444444
```

```
avg_penalties_mediumdefense= df.loc[df['defensive_work_rate'] == 1, 'penalties'].copy()
```

```
avg_penalties_mediumdefense.mean()
```

```
53.39736025556761
```

```
df.to_csv('C:\\Users\\karthik nandiraju\\Documents\\ML\\soccer.csv')
```

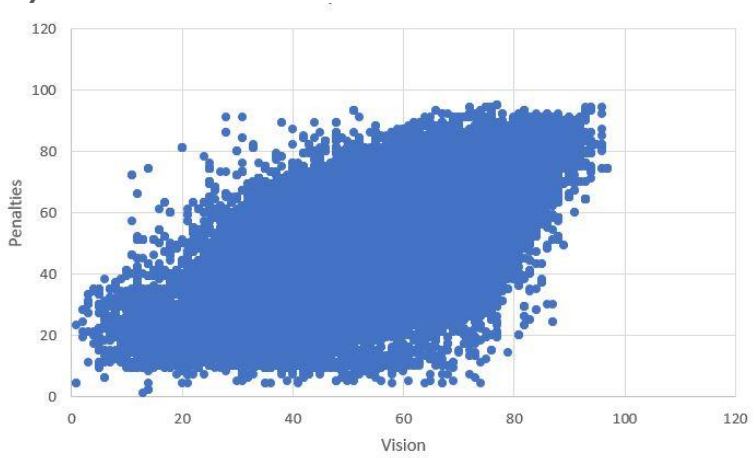
```
df['vision'].corr(df['penalties'])
```

```
0.66580169130749467
```

I could not attach my notebook.

Findings

3. I learnt that players with high vision associate with more penalties than players with low vision. The correlation between vision and penalties is fairly positive(0.665).



Acknowledgements

I had no one to give me feedback.

References

www.google.com

www.Kaggle.com