

# European Soccer Database

Mestrado em Modelagem Matemática  
Disciplina: Modelagem e Mineração de Dados

Renato Santos Aranha

22 de Setembro de 2016


## Agenda:

- Origem da base de dados;
- Perguntas formuladas;
- Abordagem e Resultados;
- Próximos passos;
- Conclusões;


## **Agenda:**

- **Origem da base de dados;**
- Perguntas formuladas;
- Abordagem e Resultados;
- Próximos passos;
- Conclusões;

<https://www.kaggle.com/datasets>



[Competitions](#)
[Datasets](#)
[Kernels](#)
[Forums](#)
[Jobs](#)




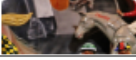


# Datasets

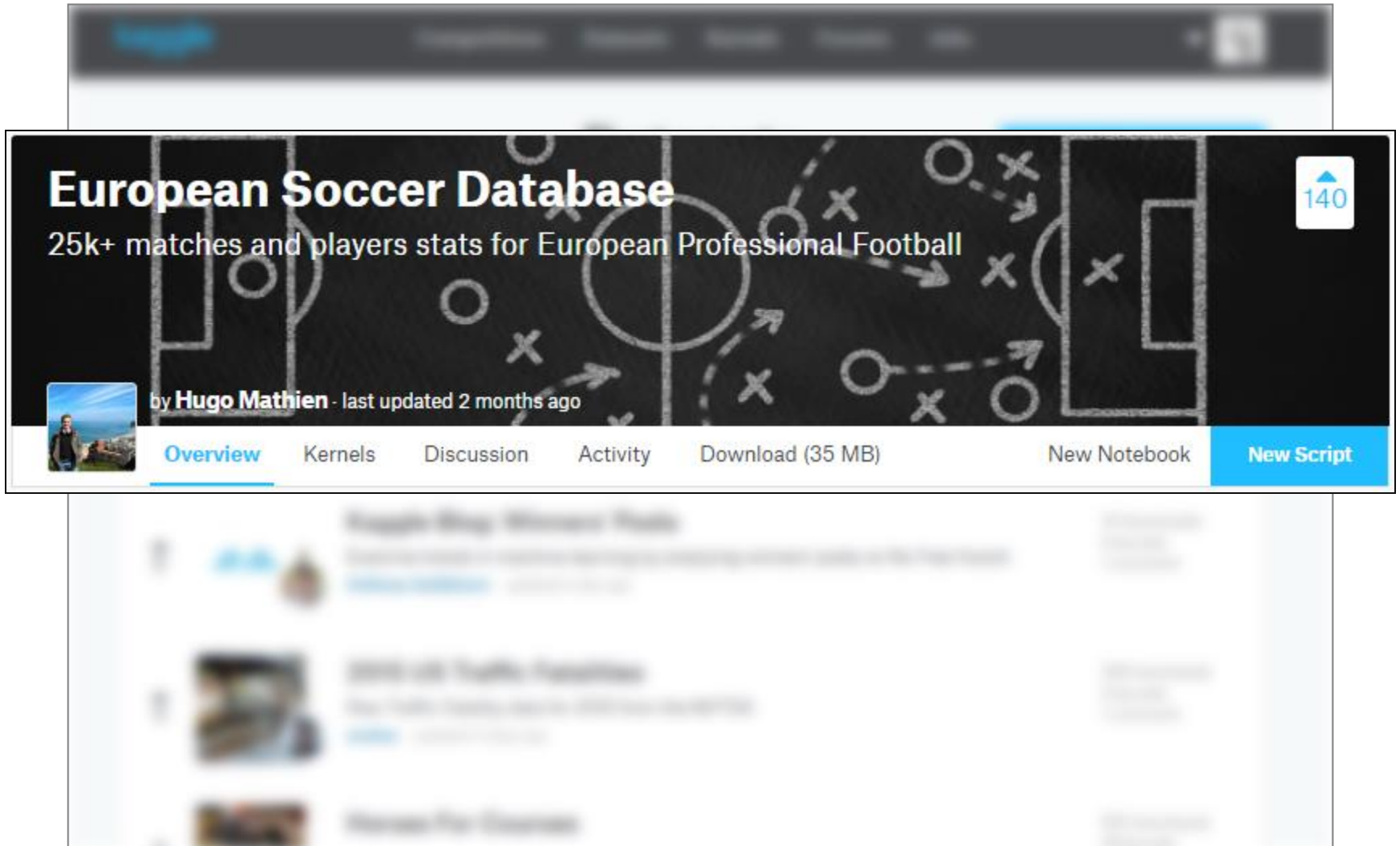
New Dataset

85 featured datasets
Sort By Hotness

Featured All Mine Upvoted

12		<b>Hospital Charges for Inpatients</b> How inpatient hospital charges can differ among different providers in the US <a href="#">Pranay Aryal</a> · updated 3 days ago	335 downloads 14 kernels 5 comments
6		<b>Kaggle Blog: Winners' Posts</b> Examine trends in machine learning by analyzing winners' posts on No Free Hunch <a href="#">Anthony Goldbloom</a> · updated a day ago	34 downloads 5 kernels 1 comment
9		<b>2015 US Traffic Fatalities</b> Raw Traffic Fatality data for 2015 from the NHTSA <a href="#">anokas</a> · updated 4 days ago	299 downloads 2 kernels 1 comment
		<b>Horses For Courses</b>	925 downloads 28 kernels

<https://www.kaggle.com/datasets>



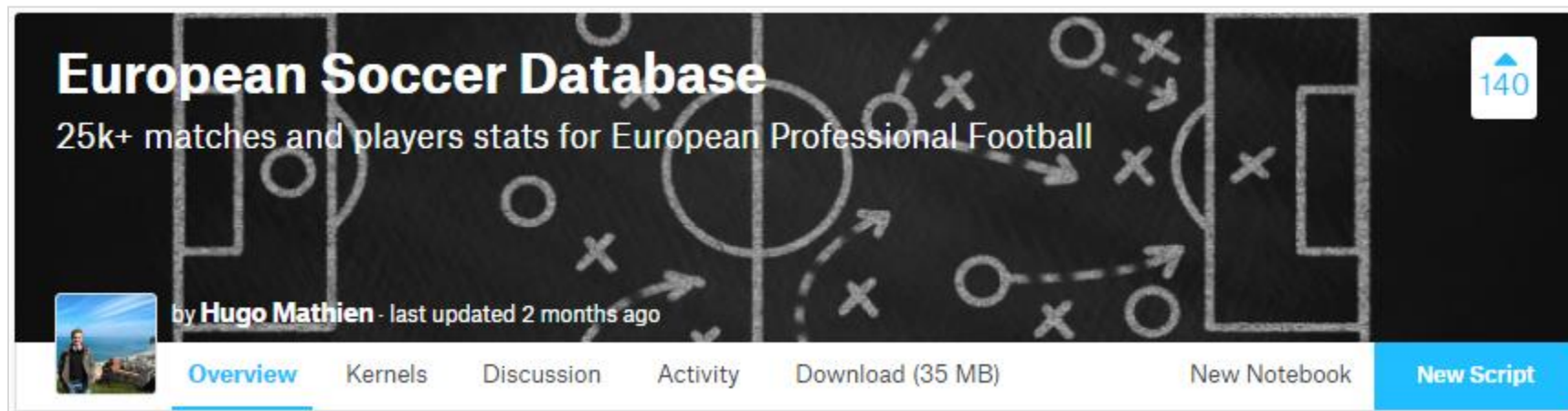
The screenshot shows the Kaggle dataset page for the "European Soccer Database". The header features a soccer field diagram with player positions and movement arrows. The title "European Soccer Database" is prominently displayed, followed by the description "25k+ matches and players stats for European Professional Football". A badge in the top right corner indicates 140 upvotes. Below the title, it states "by Hugo Mathien - last updated 2 months ago". A navigation bar includes links for "Overview", "Kernels", "Discussion", "Activity", "Download (35 MB)", "New Notebook", and "New Script". The "Overview" link is currently selected. The main content area below the navigation bar is blurred, showing a list of related datasets or kernels.

## European Soccer Database

25k+ matches and players stats for European Professional Football

by **Hugo Mathien** - last updated 2 months ago

[Overview](#) [Kernels](#) [Discussion](#) [Activity](#) [Download \(35 MB\)](#) [New Notebook](#) [New Script](#)



**European Soccer Database**  
 25k+ matches and players stats for European Professional Football

by **Hugo Mathien** - last updated 2 months ago

Overview Kernels Discussion Activity Download (35 MB) New Notebook **New Script**

140

## The ultimate Soccer database for data analysis and machine learning

### What you get:

- +25,000 matches
- +10,000 players
- 11 European Countries with their lead championship
- Seasons 2008 to 2016
- Player's attributes sourced from EA Sports' FIFA video game series, including the weekly updates
- Team line up with squad formation (X, Y coordinates)
- Betting odds from up to 10 providers
- Detailed match events (goal types, possession, corner, cross, fouls, cards etc...) for +10,000 matches
- <http://football-data.mx-api.enetscores.com/> : scores, lineup, team formation and events
- <http://www.football-data.co.uk/> : betting odds
- <http://sofifa.com/> : player attributes from EA Sports FIFA games. *FIFA series and all FIFA assets property of EA Sports.*

# European Soccer Database

25k+ matches and players stats for European Professional Football

by **Hugo Mathien** - last updated 2 months ago

140

[Overview](#)
[Kernels](#)
[Discussion](#)
[Activity](#)
[Download \(35 MB\)](#)
[New Notebook](#)
[New Script](#)

database.sqlite (35.13 MB)

## Tables:

- **Country:** 11 European countries
- **League:** 11 Leagues, one per country
- **Match:** Home and away teams/goals, lineups with XY coordinates, betting odds. This version of the database does not include the detailed match events
- **Player:** ids and a few general stats.
- **Player\_Stats:** the features from EA Sport's FIFA football video games. date\_stat indicates as of when the set of features was made available
- **Team:** teams id and names

## Foreign Keys

Country and League are referenced in the Match table by their primary key (*id*). Matches and Players are linked to other tables by foreign keys *\*match\_api\_id\** and *\*player\_api\_id\**. These keys are different from the table's primary keys (*id*). I have deliberately made the joint on these *\*api\_id\** to be consistent with the original data sources

## Agenda:

- Origem da base de dados;
- **Perguntas formuladas;**
- Abordagem e Resultados;
- Próximos passos;
- Conclusões;



- 1) Qual é o percentual de partidas vencidas pelo time mandante?
- 2) Como é a distribuição de partidas por gol de diferença (e por país)?
- 3) Qual é o percentual de partidas que terminam empatadas (e por país)?
- 4) Jogadores de mesma posição possuem semelhança nos scouts?
- 5) Como ocorreu a evolução anual (overall rating) dos melhores jogadores?

## Agenda:

- Origem da base de dados;
- Perguntas formuladas;
- **Abordagem e Resultados;**
- Próximos passos;
- Conclusões;

- 1) Qual é o percentual de partidas vencidas pelo time mandante?
- 2) Como é a distribuição de partidas por gol de diferença (e por país)?
- 3) Qual é o percentual de partidas que terminam empatadas (e por país)?
- 4) Jogadores de mesma posição possuem semelhança nos scouts?
- 5) Como ocorreu a evolução anual (overall rating) dos melhores jogadores?

## 1) Qual é o percentual de partidas vencidas pelo time mandante?

Feature engineering:  
criação de campo

Feature engineering:  
Criação de data frame  
específico para vitórias  
de mandantes;

Feature engineering:  
Criação de data frame  
específico para vitórias  
de visitantes;

```
In [135]: match['dif_gol'] = pd.Series(match['home_team_goal'] - match['away_team_goal'])

match_casa = match[match['dif_gol']>0][['país', 'dif_gol', 'match_api_id']]
match_fora = match[match['dif_gol']<0][['país', 'dif_gol', 'match_api_id']]

print ()
print ()
print (round(100*(match_casa['match_api_id'].count())/match.shape[0],1), "% das partidas são vencidas pelo time da casa.")
print ()
print (round(100*(match_fora['match_api_id'].count())/match.shape[0],1), "% das partidas são vencidas pelo time visitante.")
print ()
partidas_por_dif_de_gols = match.groupby('dif_gol').count()['match_api_id']
print (round(100*partidas_por_dif_de_gols[0]/match.shape[0],1), "% são empates." )
print ()
print ()
```

45.9 % das partidas são vencidas pelo time da casa.

28.7 % das partidas são vencidas pelo time visitante.

25.4 % são empates.

- 1) Qual é o percentual de partidas vencidas pelo time mandante?
- 2) Como é a distribuição de partidas por gol de diferença (e por país)?
- 3) Qual é o percentual de partidas que terminam empatadas (e por país)?
- 4) Jogadores de mesma posição possuem semelhança nos scouts?
- 5) Como ocorreu a evolução anual (overall rating) dos melhores jogadores?

## 2) Como é a distribuição de partidas por gol de diferença (e por país)?

\* Como é a distribuição de partidas por gol de diferença?

```
In [134]: match['dif_gol'] = abs(pd.Series(match['home_team_goal'] - match['away_team_goal'])) #criando coluna com a diferença de gols

jogos_dif_de_gols = match.groupby(['dif_gol']).count()['match_api_id']
jogos_dif_de_gols.plot(kind='bar',figsize=(15,6),title='QUANTIDADE DE PARTIDAS POR DIFERENÇA DE GOLS', color = 'g')

empate = jogos_dif_de_gols[0];
um_gol_de_dif = jogos_dif_de_gols[1];
dois_gols_de_dif = jogos_dif_de_gols[2];
print ()
print (round(100*(empate)/match.shape[0],1), "% das partidas terminam empatadas.")
print ("(",empate,"em ",match.shape[0],"jogos).")
print ()
print (round(100*(um_gol_de_dif)/match.shape[0],1), "% das partidas terminam com diferença de UM gol.")
print ("(",um_gol_de_dif,"em ",match.shape[0],"jogos).")
print ()
print (round(100*(dois_gols_de_dif)/match.shape[0],1), "% das partidas terminam com diferença de DOIS gols.")
print ("(",dois_gols_de_dif,"em ",match.shape[0],"jogos).")
```

25.4 % das partidas terminam empatadas.  
( 6590 em 25945 jogos).

37.0 % das partidas terminam com diferença de UM gol.  
( 9587 em 25945 jogos).

22.1 % das partidas terminam com diferença de DOIS gols.  
( 5728 em 25945 jogos).

## 2) Como é a distribuição de partidas por gol de diferença (e por país)?

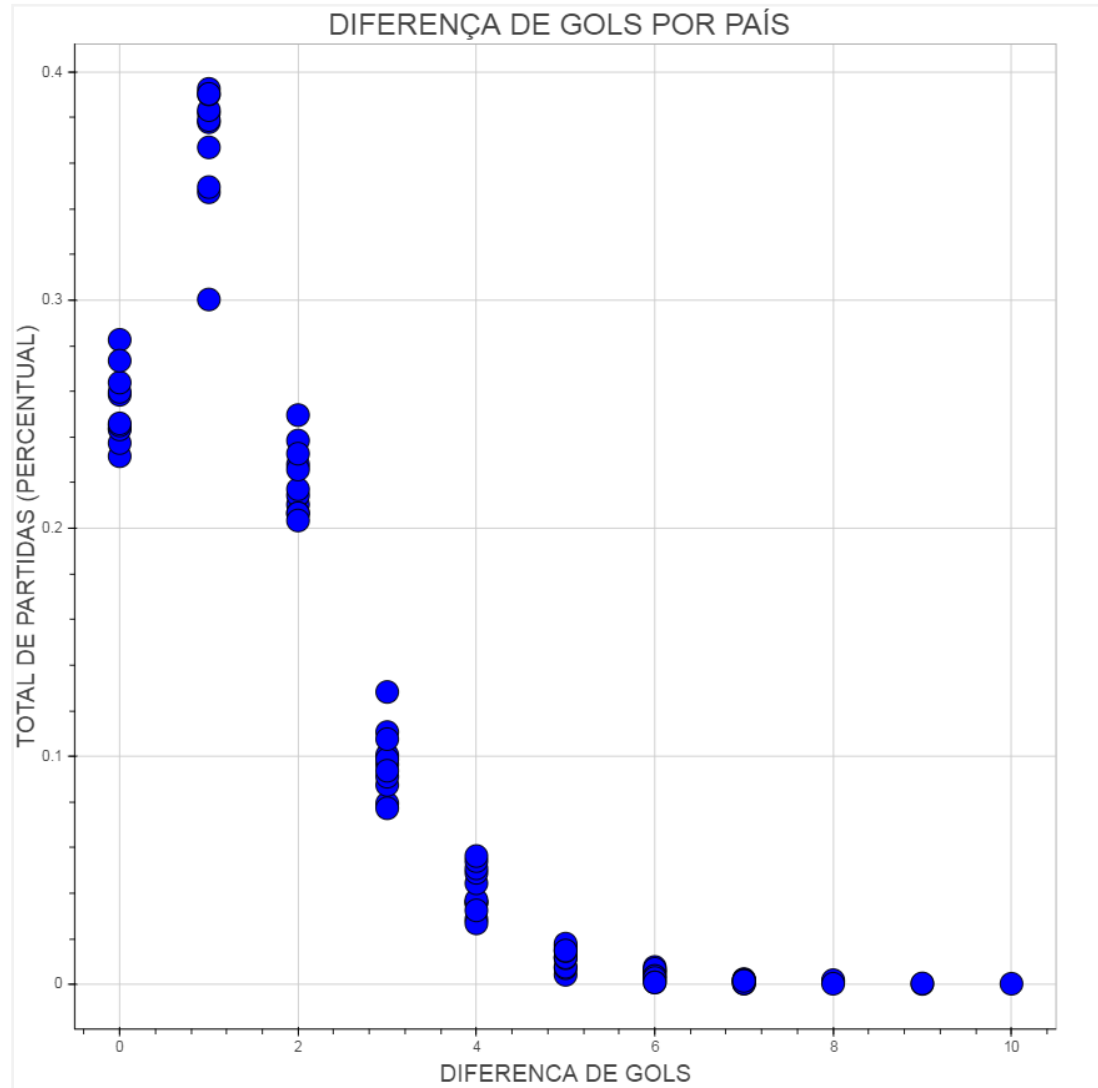
25.4 % das partidas terminam empatadas.  
 ( 6590 em 25945 jogos).

37.0 % das partidas terminam com diferença de UM gol.  
 ( 9587 em 25945 jogos).

22.1 % das partidas terminam com diferença de DOIS gols.  
 ( 5728 em 25945 jogos).



2) Como é a distribuição de partidas por gol de diferença (e por país)?

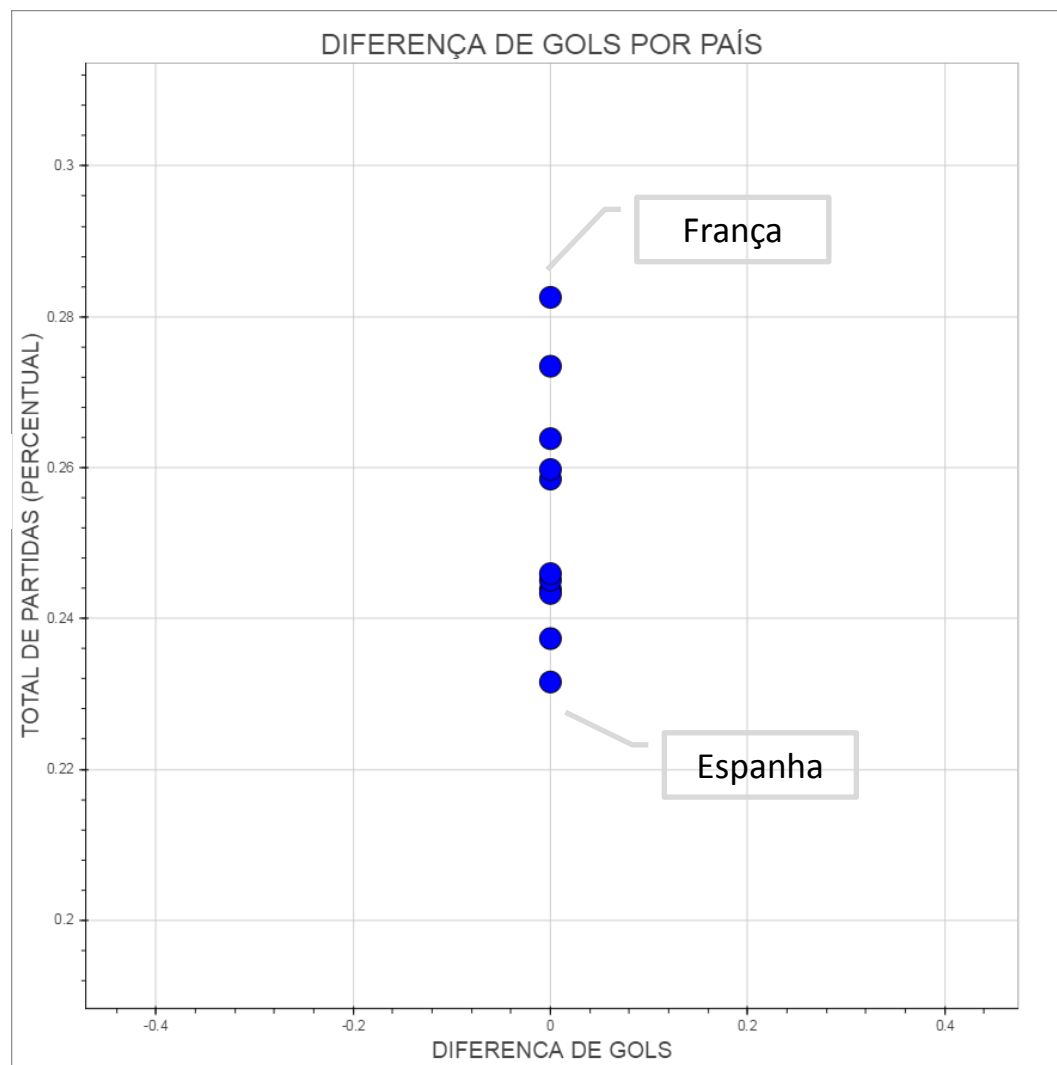




- 1) Qual é o percentual de partidas vencidas pelo time mandante?
- 2) Como é a distribuição de partidas por gol de diferença (e por país)?
- 3) Qual é o percentual de partidas que terminam empatadas (e por país)?
- 4) Jogadores de mesma posição possuem semelhança nos scouts?
- 5) Como ocorreu a evolução anual (overall rating) dos melhores jogadores?

### 3) Qual é o percentual de partidas que terminam empatadas (e por país)?

25.4 % das partidas terminam empatadas.  
 ( 6590 em 25945 jogos).

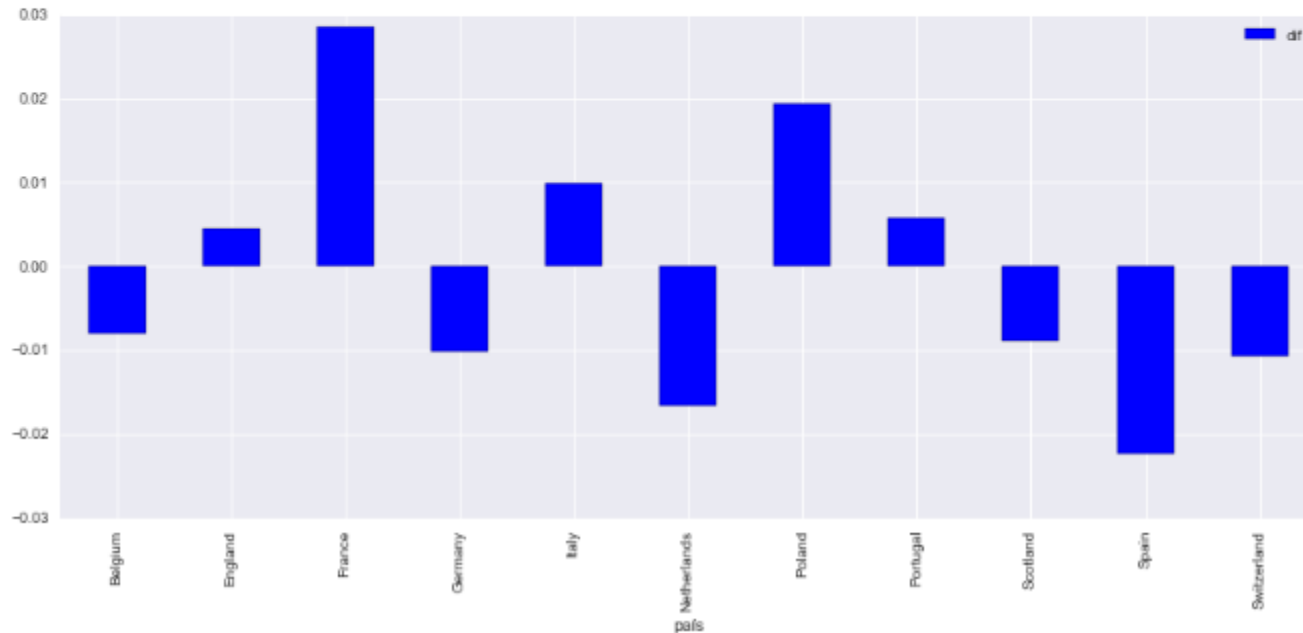


### 3) Qual é o percentual de partidas que terminam empatadas (e por país)?

```
In [34]: novo['percent']=novo['match_api_id_x']/novo['match_api_id_y']
graf = novo[novo['dif_gol']==0][['país','percent']]
graf['dif'] = graf['percent']-0.254 #0,254 é o percentual total de empates em todas as ligas (aqui estou pegando os desvios)
graf.plot(kind='bar',figsize=(15,6), color = 'b', x='país',y='dif')
print()
print("Países com MAIOR incidência de empates: France e Poland")
print()
print("Países com MENOR incidência de empates: Spain e Netherlands")
print()
```

Países com MAIOR incidência de empates: France e Poland

Países com MENOR incidência de empates: Spain e Netherlands



- 1) Qual é o percentual de partidas vencidas pelo time mandante?
- 2) Como é a distribuição de partidas por gol de diferença (e por país)?
- 3) Qual é o percentual de partidas que terminam empatadas (e por país)?
- 4) Jogadores de mesma posição possuem semelhança nos scouts?
- 5) Como ocorreu a evolução anual (overall rating) dos melhores jogadores?

## 4) Jogadores de mesma posição possuem semelhança nos scouts?

### t-distributed stochastic neighbor embedding

From Wikipedia, the free encyclopedia

**t-distributed stochastic neighbor embedding (t-SNE)** is a [machine learning](#) algorithm for [dimensionality reduction](#) developed by [Laurens van der Maaten](#)<sup>[1]</sup> and [Geoffrey Hinton](#).<sup>[1]</sup> It is a [nonlinear dimensionality reduction](#) technique that is particularly well-suited for embedding high-dimensional data into a space of two or three dimensions, which can then be visualized in a [scatter plot](#). Specifically, it models each high-dimensional object by a two- or three-dimensional point in such a way that similar objects are modeled by nearby points and dissimilar objects are modeled by distant points.

The t-SNE algorithm comprises two main stages. First, t-SNE constructs a [probability distribution](#) over pairs of high-dimensional objects in such a way that similar objects have a high probability of being picked, whilst dissimilar points have an extremely small probability of being picked. Second, t-SNE defines a similar probability distribution over the points in the low-dimensional map, and it minimizes the [Kullback–Leibler divergence](#) between the two distributions with respect to the locations of the points in the map. Note that whilst the original algorithm uses the [Euclidean distance](#) between objects as the base of its similarity metric, this should be changed as appropriate.


t-SNE has been used in a wide range of applications, including [computer security research](#),<sup>[2]</sup> [music analysis](#),<sup>[3]</sup> [cancer research](#),<sup>[4]</sup> [bioinformatics](#),<sup>[5]</sup> and [biomedical signal processing](#).<sup>[6]</sup>

#### Details [\[ edit \]](#)

Given a set of  $N$  high-dimensional objects  $\mathbf{x}_1, \dots, \mathbf{x}_N$ , t-SNE first computes probabilities  $p_{ij}$  that are proportional to the similarity of objects  $\mathbf{x}_i$  and  $\mathbf{x}_j$ , as follows:

$$p_{j|i} = \frac{\exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma_i^2)}{\sum_{k \neq i} \exp(-\|\mathbf{x}_i - \mathbf{x}_k\|^2 / 2\sigma_i^2)},$$

### Machine learning and data mining



Problems	<a href="#">[show]</a>
Supervised learning (classification • regression)	<a href="#">[show]</a>
Clustering	<a href="#">[show]</a>
Dimensionality reduction	<a href="#">[show]</a>
Structured prediction	<a href="#">[show]</a>
Anomaly detection	<a href="#">[show]</a>
Neural nets	<a href="#">[show]</a>
Reinforcement Learning	<a href="#">[show]</a>
Theory	<a href="#">[show]</a>
Machine learning venues	<a href="#">[show]</a>
<a href="#">Machine learning portal</a>	
V · T · E	

#### 4) Jogadores de mesma posição possuem semelhança nos scouts?

### sklearn.manifold.TSNE

```
class sklearn.manifold. TSNE (n_components=2, perplexity=30.0, early_exaggeration=4.0, learning_rate=1000.0,
n_iter=1000, n_iter_without_progress=30, min_grad_norm=1e-07, metric='euclidean', init='random', verbose=0,
random_state=None, method='barnes_hut', angle=0.5)
```

[\[source\]](#)

t-distributed Stochastic Neighbor Embedding.

t-SNE [1] is a tool to visualize high-dimensional data. It converts similarities between data points to joint probabilities and tries to minimize the Kullback-Leibler divergence between the joint probabilities of the low-dimensional embedding and the high-dimensional data. t-SNE has a cost function that is not convex, i.e. with different initializations we can get different results.

It is highly recommended to use another dimensionality reduction method (e.g. PCA for dense data or TruncatedSVD for sparse data) to reduce the number of dimensions to a reasonable amount (e.g. 50) if the number of features is very high. This will suppress some noise and speed up the computation of pairwise distances between samples. For more tips see Laurens van der Maaten's FAQ [2].

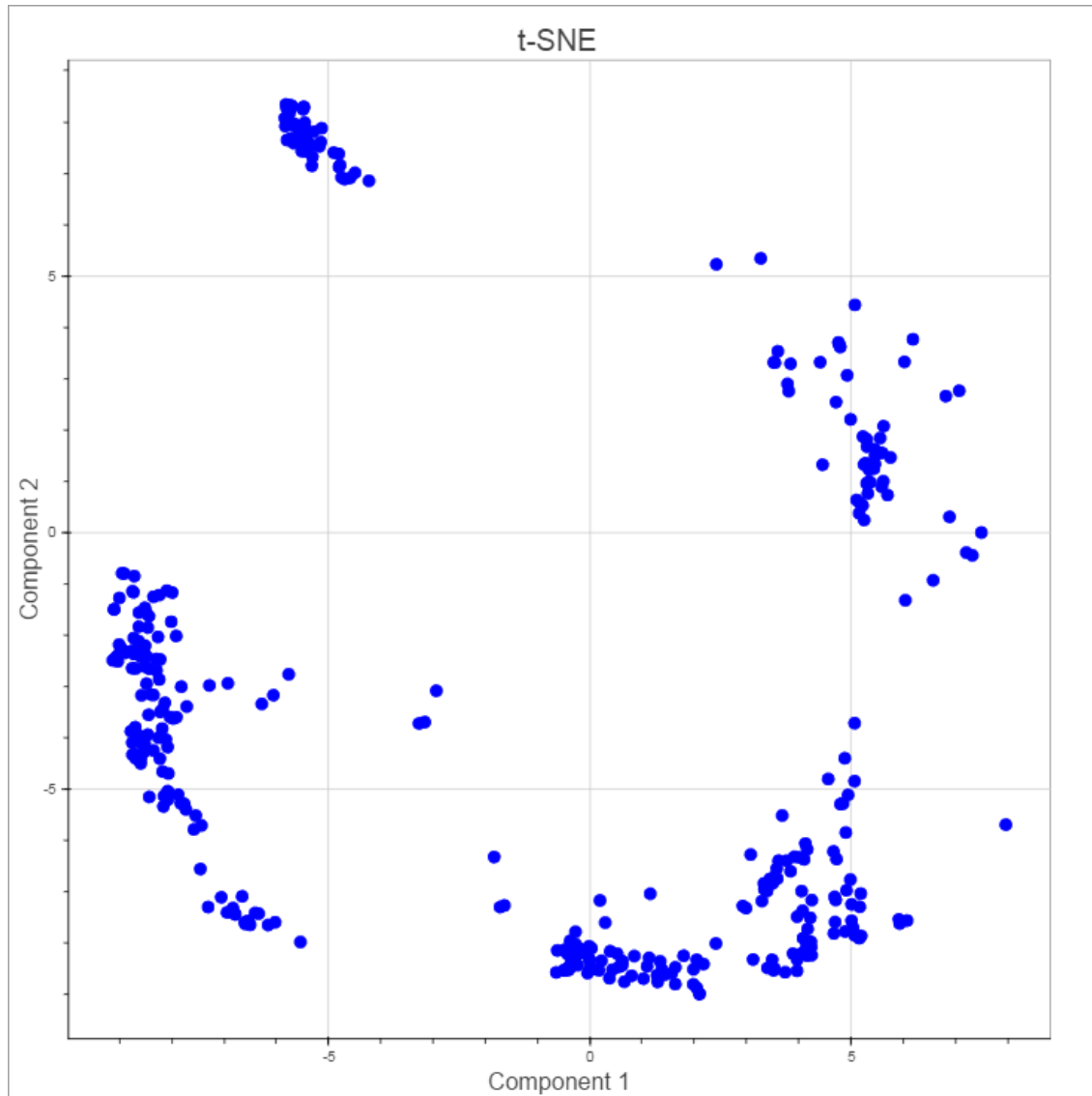
Read more in the [User Guide](#).

**Parameters:** **n\_components** : int, optional (default: 2)

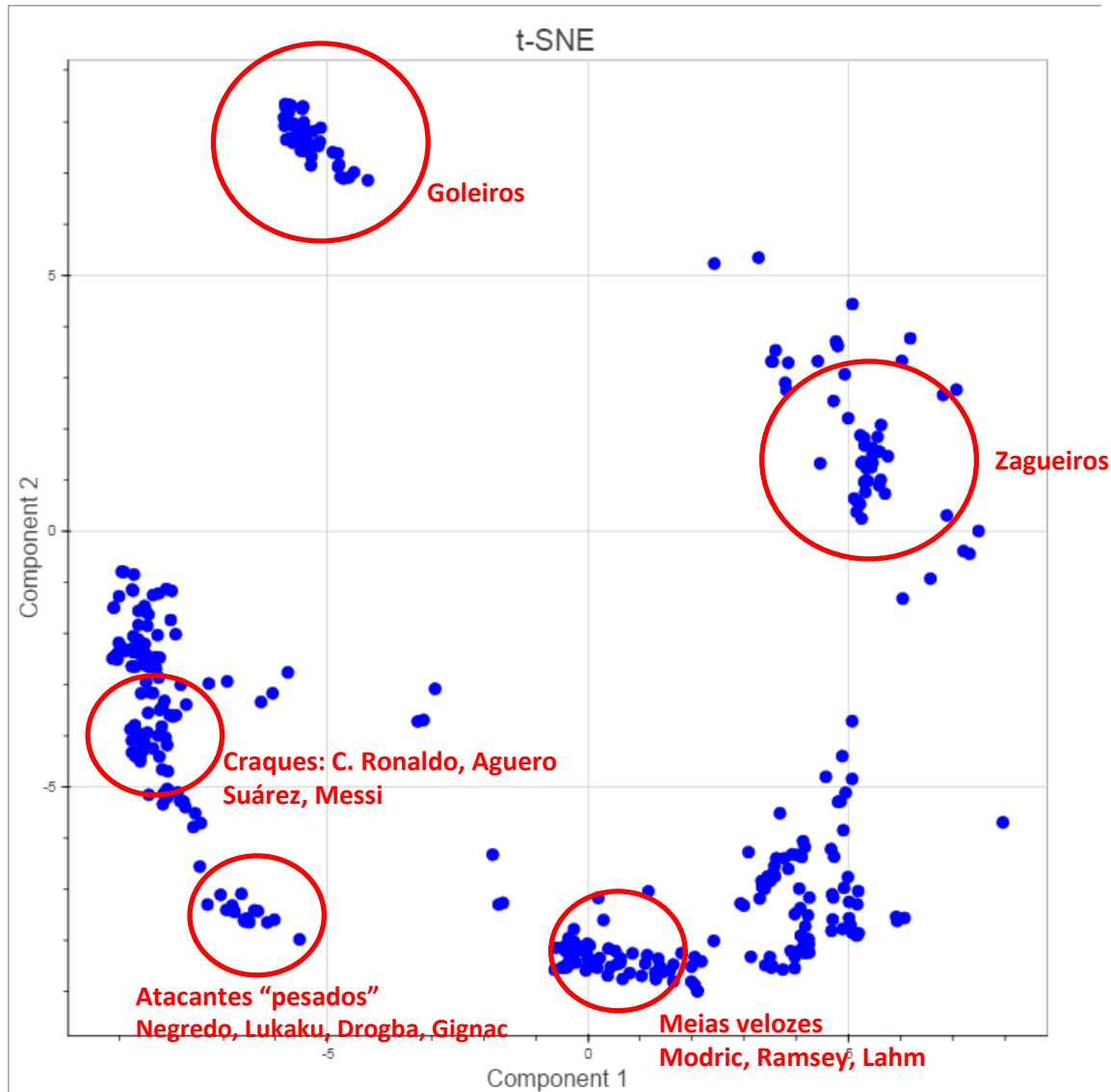
Dimension of the embedded space.

**perplexity** : float, optional (default: 30)

#### 4) Jogadores de mesma posição possuem semelhança nos scouts?



#### 4) Jogadores de mesma posição possuem semelhança nos scouts?





- 1) Qual é o percentual de partidas vencidas pelo time mandante?
- 2) Como é a distribuição de partidas por gol de diferença (e por país)?
- 3) Qual é o percentual de partidas que terminam empatadas (e por país)?
- 4) Jogadores de mesma posição possuem semelhança nos scouts?
- 5) Como ocorreu a evolução anual (overall rating) dos melhores jogadores?

## 5) Como ocorreu a evolução anual (overall rating) dos melhores jogadores?

```
In [148]: players = players.sort_values('overall_rating', ascending=False)
best_players = players[['player_api_id', 'player_name']].head(15)
ids = tuple(best_players.player_api_id.unique())

query = '''SELECT player_api_id, date_stat, overall_rating, potential
          FROM Player_Stats WHERE player_api_id in %s''' % (ids,)

evolution = pd.read_sql(query, conn)
evolution = pd.merge(evolution, best_players)
evolution['year'] = evolution.date_stat.str[:4].apply(int)
evolution = evolution.groupby(['year', 'player_api_id', 'player_name']).overall_rating.mean()
evolution = evolution.reset_index()

#evolution.sort_values('overall_rating', ascending=False).head(10)
best_players
```

Out[148]:

	player_api_id	player_name
101978	30981	Lionel Messi
33156	30893	Cristiano Ronaldo
130834	19533	Neymar
105468	40636	Luis Suarez
108509	27299	Manuel Neuer
16383	30834	Arjen Robben
182801	35724	Zlatan Ibrahimovic
121856	36378	Mesut Oezil
12717	30955	Andres Iniesta
47010	107417	Eden Hazard
149822	93447	Robert Lewandowski
160061	37412	Sergio Aguero
168976	80562	Thiago Silva
105683	31097	Luka Modric
39006	182917	David De Gea

## 5) Como ocorreu a evolução anual (overall rating) dos melhores jogadores?

```
In [152]: players = players.sort_values('overall_rating', ascending=False)
best_players = players[['player_api_id', 'player_name']].head(15)
ids = tuple(best_players.player_api_id.unique())

query = '''SELECT player_api_id, date_stat, overall_rating, potential
          FROM Player_Stats WHERE player_api_id in %s''' % (ids,)

evolution = pd.read_sql(query, conn)
evolution = pd.merge(evolution, best_players)
evolution['year'] = evolution.date_stat.str[:4].apply(int)
evolution = evolution.groupby(['year', 'player_api_id', 'player_name']).overall_rating.mean()
evolution = evolution.reset_index()

evolution.sort_values('overall_rating', ascending=False).head(15)
#best_players
```

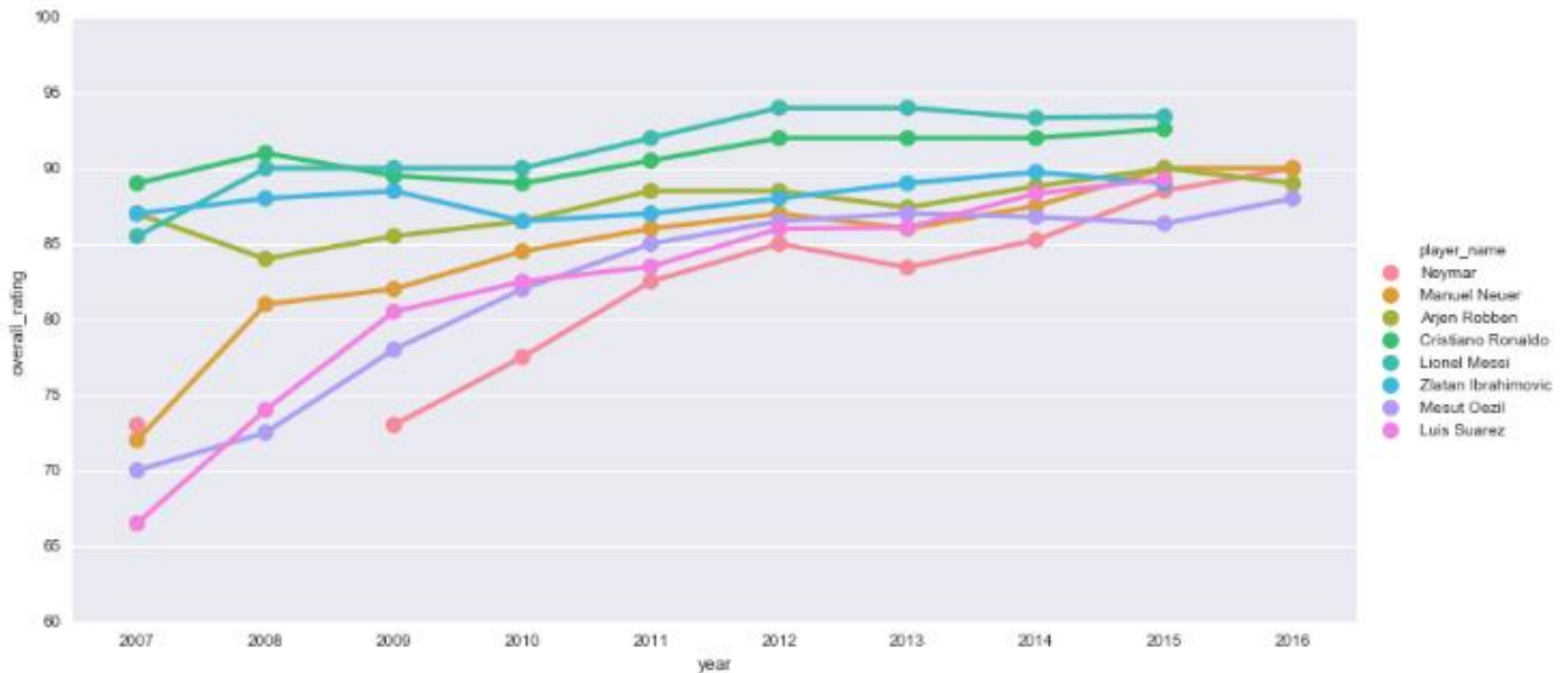
Out[152]:

	year	player_api_id	player_name	overall_rating
93	2013	30981	Lionel Messi	94.000000
78	2012	30981	Lionel Messi	94.000000
123	2015	30981	Lionel Messi	93.428571
108	2014	30981	Lionel Messi	93.333333
121	2015	30893	Cristiano Ronaldo	92.600000
91	2013	30893	Cristiano Ronaldo	92.000000
106	2014	30893	Cristiano Ronaldo	92.000000
63	2011	30981	Lionel Messi	92.000000
76	2012	30893	Cristiano Ronaldo	92.000000
17	2008	30893	Cristiano Ronaldo	91.000000
61	2011	30893	Cristiano Ronaldo	90.500000
77	2012	30955	Andres Iniesta	90.500000
48	2010	30981	Lionel Messi	90.000000
120	2015	30834	Arjen Robben	90.000000
19	2008	30981	Lionel Messi	90.000000

## 5) Como ocorreu a evolução anual (overall rating) dos melhores jogadores?

Evolução dos melhores jogadores:

```
In [153]: a = sns.factorplot(data=evolution[evolution.player_api_id.isin(ids[0:8])], x='year',
                             y='overall_rating', hue='player_name', size=6, aspect=2, kind='point', orient='v', legend_out=True)
```



## Agenda:

- Origem da base de dados;
- Perguntas formuladas;
- Abordagem e Resultados;
- **Próximos passos;**
- Conclusões;

- 1) Qual é o percentual de partidas vencidas pelo time mandante?
- 2) Como é a distribuição de partidas por gol de diferença (e por país)?
- 3) Qual é o percentual de partidas que terminam empatadas (e por país)?
- 4) Jogadores de mesma posição possuem semelhança nos scouts?
  - **Visualização de scouts por radar;**
  - **Classificador para definir a liga onde o jogador atua;**
- 5) Como ocorreu a evolução anual (overall rating) dos melhores jogadores?
  - **E dos piores jogadores?**
  - **Em qual idade ocorre o auge?**
- 6) **Análise preditiva do resultados das partidas, por país;**
  - **SVM parece ser bem indicado.**

## Agenda:

- Origem da base de dados;
- Perguntas formuladas;
- Abordagem e Resultados;
- Próximos passos;
- **Conclusões;**

- 1) Qual é o percentual de partidas vencidas pelo time mandante?
  - **Aprox. 46%. O interessante é que essa média é igual para todas as ligas da Europa.**
- 2) Como é a distribuição de partidas por gol de diferença (e por país)?
  - Aprox. 37% da partidas terminam com 1 gol de diferença. Outlier: liga Holandesa, onde esse percentual é de aprox. 30%.
- 3) Qual é o percentual de partidas que terminam empatadas (e por país)?
  - Aprox. 25% das partidas terminam empatadas.
  - Percentualmente, a liga francesa é a que tem mais empates (28%), e a liga espanhola (BBVA) é a que tem menos (22%).
- 4) Jogadores de mesma posição possuem semelhança nos scouts?
  - A redução de dimensionalidade via t-SNE (2 eixos) indica que scouts semelhantes têm correlação com a posição do atleta.
- 5) Como ocorreu a evolução anual (overall rating) dos melhores jogadores?
  - Através de observação superficial do gráfico, vemos que os top jogadores demoram 4~5 anos para atingirem o auge.



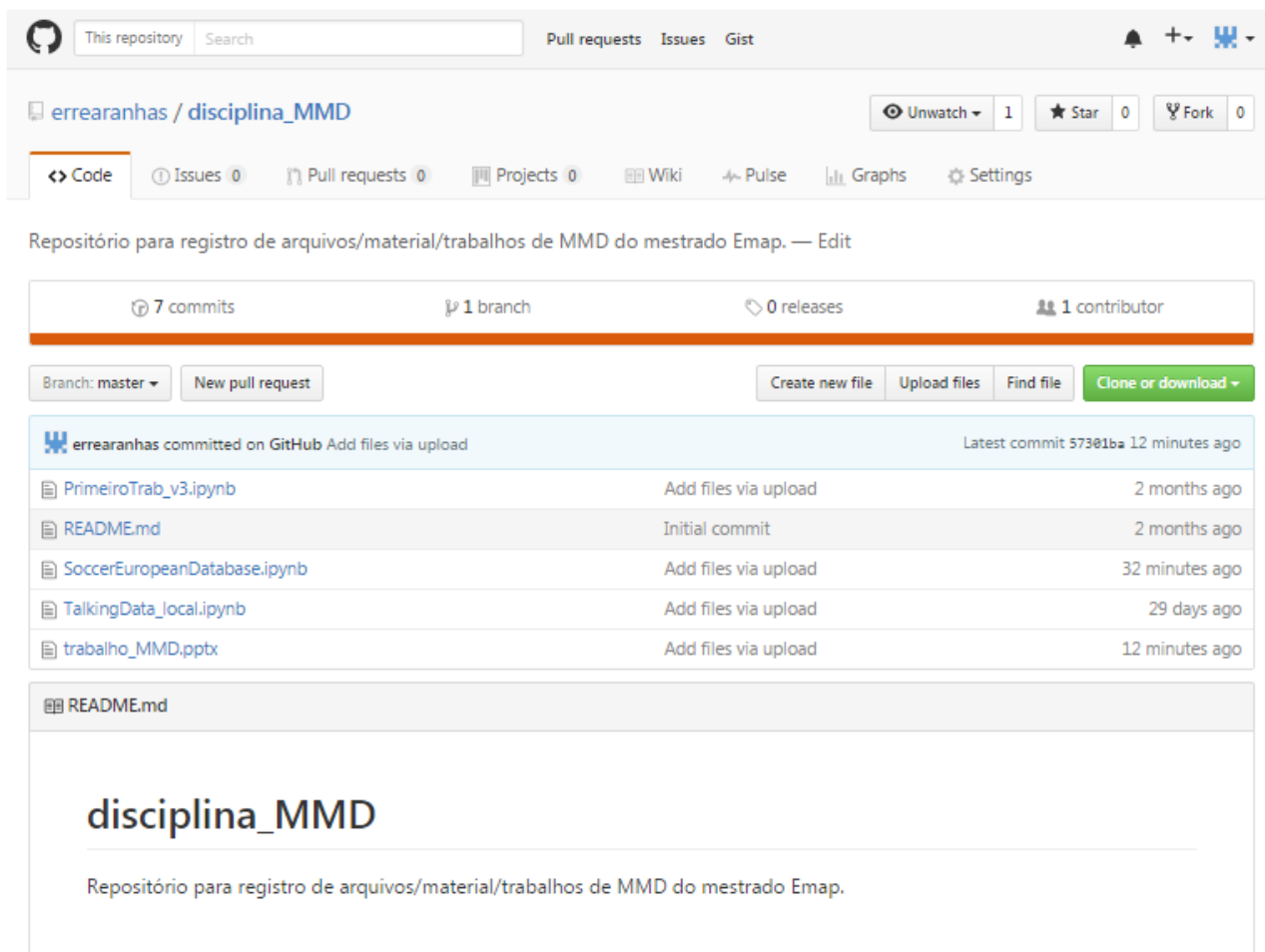
- 1) Qual é o percentual de partidas vencidas pelo time mandante?
  - **Aprox. 46%. O interessante é que essa média é igual para todas as ligas da Europa.**
- 2) Como é a distribuição de partidas por gol de diferença (e por país)?
  - **Aprox. 37% da partidas terminam com 1 gol de diferença. Outlier: liga Holandesa, onde esse percentual é de aprox. 30%.**
- 3) Qual é o percentual de partidas que terminam empatadas (e por país)?
  - **Aprox. 25% das partidas terminam empatadas.**
  - **Percentualmente, a liga francesa é a que tem mais empates (28%), e a liga espanhola (BBVA) é a que tem menos (22%).**
- 4) Jogadores de mesma posição possuem semelhança nos scouts?
  - **A redução de dimensionalidade via t-SNE (2 eixos) indica que scouts semelhantes têm correlação com a posição do atleta.**
- 5) Como ocorreu a evolução anual (overall rating) dos melhores jogadores?
  - **Através de observação superficial do gráfico, vemos que os top jogadores demoram 4~5 anos para atingirem o auge.**

- 1) Qual é o percentual de partidas vencidas pelo time mandante?
  - **Aprox. 46%. O interessante é que essa média é igual para todas as ligas da Europa.**
- 2) Como é a distribuição de partidas por gol de diferença (e por país)?
  - **Aprox. 37% das partidas terminam com 1 gol de diferença. Outlier: liga Holandesa, onde esse percentual é de aprox. 30%.**
- 3) Qual é o percentual de partidas que terminam empatadas (e por país)?
  - **Aprox. 25% das partidas terminam empatadas.**
  - **Percentualmente, a liga francesa é a que tem mais empates (28%), e a liga espanhola (BBVA) é a que tem menos (22%).**
- 4) Jogadores de mesma posição possuem semelhança nos scouts?
  - **A redução de dimensionalidade via t-SNE (2 eixos) indica que scouts semelhantes têm correlação com a posição do atleta.**
- 5) Como ocorreu a evolução anual (overall rating) dos melhores jogadores?
  - **Através de observação superficial do gráfico, vemos que os top jogadores demoram 4~5 anos para atingirem o auge.**

- 1) Qual é o percentual de partidas vencidas pelo time mandante?
  - **Aprox. 46%. O interessante é que essa média é igual para todas as ligas da Europa.**
- 2) Como é a distribuição de partidas por gol de diferença (e por país)?
  - **Aprox. 37% da partidas terminam com 1 gol de diferença. Outlier: liga Holandesa, onde esse percentual é de aprox. 30%.**
- 3) Qual é o percentual de partidas que terminam empatadas (e por país)?
  - **Aprox. 25% das partidas terminam empatadas.**
  - **Percentualmente, a liga francesa é a que tem mais empates (28%), e a liga espanhola (BBVA) é a que tem menos (22%).**
- 4) Jogadores de mesma posição possuem semelhança nos scouts?
  - **A redução de dimensionalidade via t-SNE (2 eixos) indica que scouts semelhantes têm correlação com a posição do atleta.**
- 5) Como ocorreu a evolução anual (overall rating) dos melhores jogadores?
  - **Através de observação superficial do gráfico, vemos que os top jogadores demoram 4~5 anos para atingirem o auge.**

- 1) Qual é o percentual de partidas vencidas pelo time mandante?
  - **Aprox. 46%. O interessante é que essa média é igual para todas as ligas da Europa.**
  
- 2) Como é a distribuição de partidas por gol de diferença (e por país)?
  - **Aprox. 37% das partidas terminam com 1 gol de diferença. Outlier: liga Holandesa, onde esse percentual é de aprox. 30%.**
  
- 3) Qual é o percentual de partidas que terminam empatadas (e por país)?
  - **Aprox. 25% das partidas terminam empatadas.**
  
  - **Percentualmente, a liga francesa é a que tem mais empates (28%), e a liga espanhola (BBVA) é a que tem menos (22%).**
  
- 4) Jogadores de mesma posição possuem semelhança nos scouts?
  - **A redução de dimensionalidade via t-SNE (2 eixos) indica que scouts semelhantes têm correlação com a posição do atleta.**
  
- 5) Como ocorreu a evolução anual (overall rating) dos melhores jogadores?
  - **Através de observação superficial do gráfico, vemos que os top jogadores demoram 4~5 anos para atingirem o auge.**

[https://github.com/errearanhas/disciplina\\_MMD](https://github.com/errearanhas/disciplina_MMD)



The screenshot shows the GitHub repository page for 'errearanhas/disciplina\_MMD'. The repository is described as 'Repositório para registro de arquivos/material/trabalhos de MMD do mestrado Emap.' and has 7 commits, 1 branch, 0 releases, and 1 contributor. The repository is on the 'master' branch. The file list shows the following files and their commit history:

File	Commit Message	Time
PrimeiroTrab_v3.ipynb	Add files via upload	2 months ago
README.md	Initial commit	2 months ago
SoccerEuropeanDatabase.ipynb	Add files via upload	32 minutes ago
TalkingData_local.ipynb	Add files via upload	29 days ago
trabalho_MMD.pptx	Add files via upload	12 minutes ago

The README.md file contains the following text:

# disciplina\_MMD

Repositório para registro de arquivos/material/trabalhos de MMD do mestrado Emap.