

ATP Players

Apresentação Semanal [4]

UC | Projeto Aplicado a Ciência de Dados I

Docentes | Diana Mendes & Sérgio Moro

Grupo 2

André Silvestre N°104532

Diogo Catarino N°104745

Francisco Gomes N°104944

Rita Matos N°104936

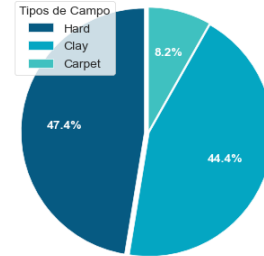
CDB1

AED

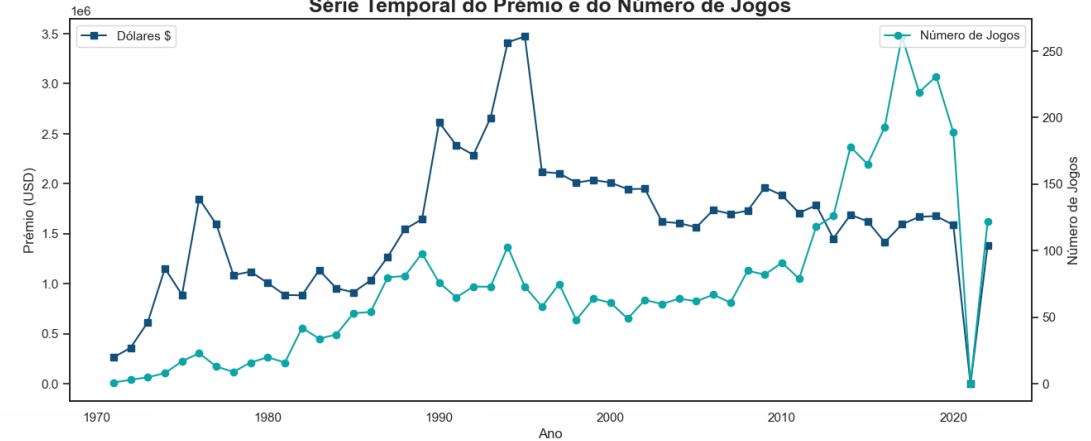
Análise Exploratória dos Dados

	n	%
Nº de Sets		
2	2623	66.20
3	1315	33.19
1	24	0.61

Distribuição dos Terrenos nos Jogos da Suécia

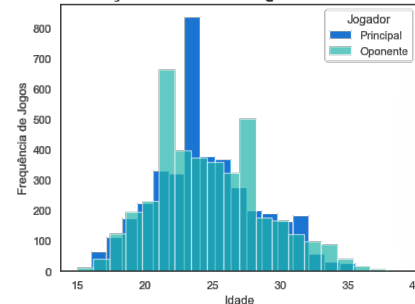


Série Temporal do Prémio e do Número de Jogos

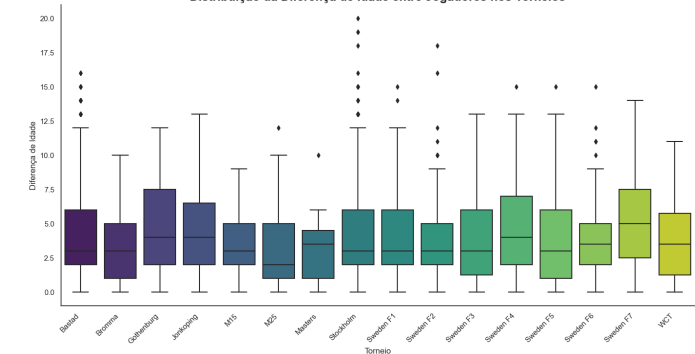


	Players (n)	Players (%)	Opponents (n)	Opponent (%)
Naturalidade Sueca				
Não	713	90.48	693	59.84
Sim	75	9.52	465	40.16
Mão Possante				
Destro	3411	86.09	3489	88.08
Canhoto	551	13.91	472	11.92
Backhand				
Two-Handed Backhand	1977	49.90	1948	49.17
Unknown Backhand	1366	34.48	1491	37.63
One-Handed Backhand	619	15.62	523	13.20

Distribuição de Idade dos Jogadores nos Torneios

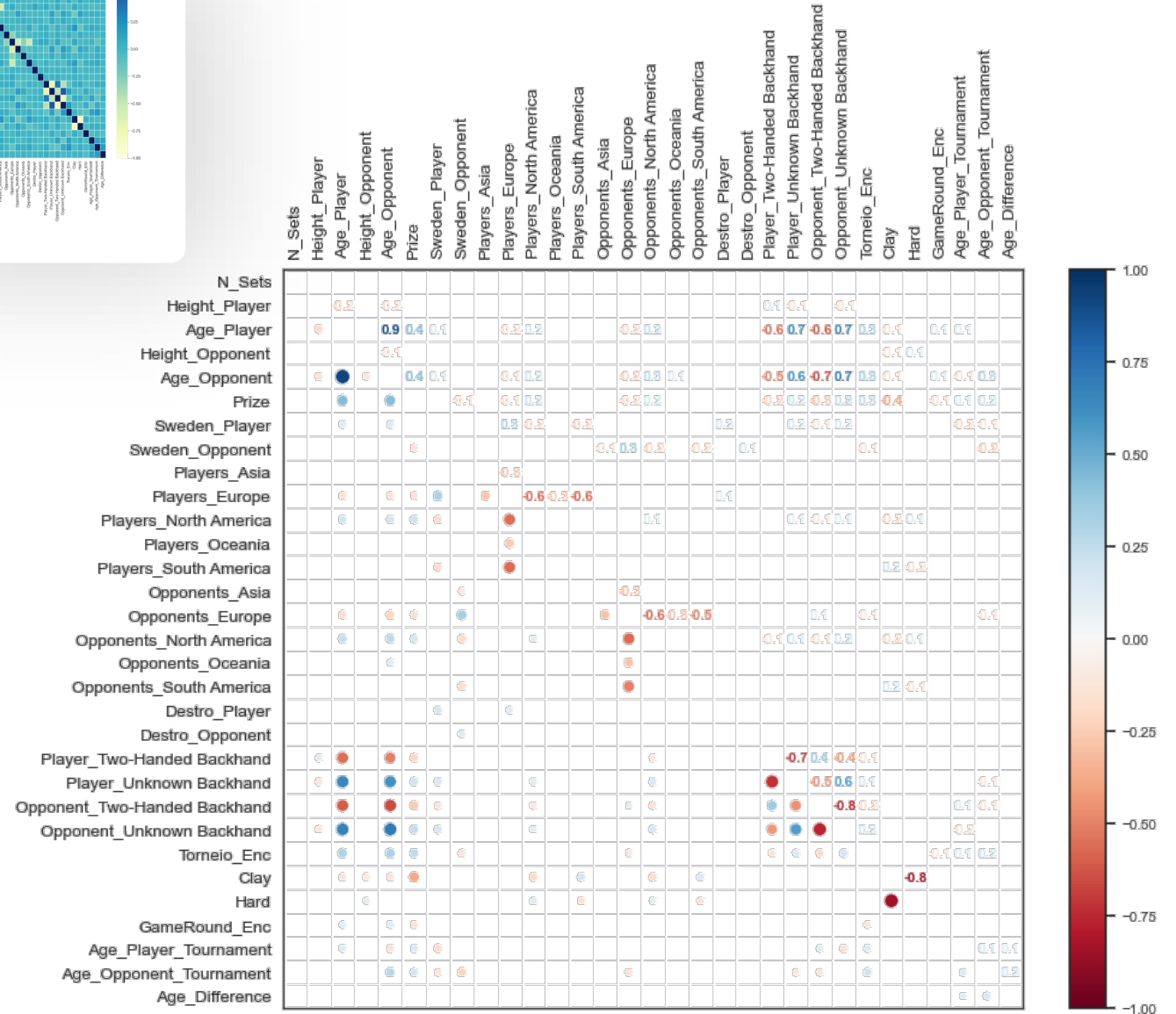
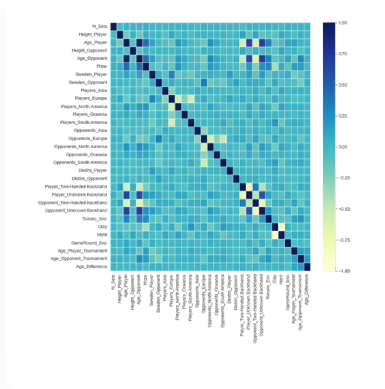


Distribuição da Diferença de Idade entre Jogadores nos Torneios



Correlações de Pearson

das Variáveis Numéricas



Variáveis Escolhidas

3973 Observações
17 Variáveis



Sweden? (Player & Opponent)

Os jogadores que jogam em ambientes familiares podem apresentar um melhor desempenho.



Diferença de Idade (Idades na altura do torneio)

As idades dos atletas podem afetar a sua capacidade física.
É possível que um jogador mais jovem tenha vantagem sobre um jogador mais velho.



Mão (Player & Opponent)

A preferência da mão dominante dos jogadores pode influenciar a dinâmica do jogo.



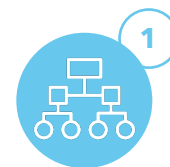
Altura (Player & Opponent)

A altura pode influenciar a habilidade e estratégia dos jogadores.



Torneio

O tipo de torneio é provável ser uma característica influenciadora do nº de sets.



Round

A ronda em que o jogo ocorre pode influenciar o nível de desempenho e motivação



Ground

O tipo de superfície pode influenciar o estilo de jogo e a estratégia.



Prize

O valor do prémio em disputa pode afetar o nível de motivação e determinação.



Modelo 1.

Treino/Teste = 90/10

KNN

K-Nearest Neighbors

Accuracy

0.57

Precision

0.52

Recall

0.57

F1-Score

0.53

AUC Score Average

0.47

Confusion matrix

True	1	0	2	0
	2	1	204	57
	3	0	110	21
		0	1 Pred	2



Modelo 2.

Treino/Teste = 90/10

Regressão Logística

Accuracy	Precision	Recall	F1-Score
0.66	0.44	0.66	0.53
AUC Score Average			
0.56			

Confusion matrix

True	1	0	2	0
	2	0	262	0
	3	0	131	0
		0	1 Pred	2



Modelo 3.

Treino/Teste = 90/10

Árvore de Decisão

Accuracy	Precision	Recall	F1-Score
0.56	0.56	0.56	0.56
AUC Score Average			
0.51			

Confusion matrix

True	1	0	1	1
	2	0	179	83
	3	1	86	44
		0	1 Pred	2



Modelo 4.

Treino/Teste = 90/10

Random Forest

Accuracy	Precision	Recall	F1-Score
0.65	0.58	0.65	0.57
AUC Score Average			
0.60			

Confusion matrix

True	1	0	2	0
	2	0	236	26
	3	0	112	19
		0	1 Pred	2

**Modelo 1.**

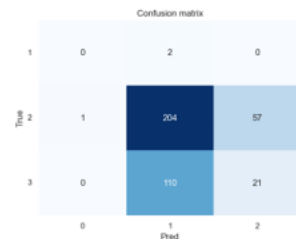
Treino/Teste = 90/10

KNN
K-Nearest Neighbors

Accuracy	Precision	Recall	F1-Score
0.57	0.52	0.57	0.53

AUC Score Average

0.47



5

**Modelo 2.**

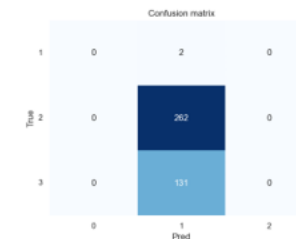
Treino/Teste = 90/10

Regressão Logística

Accuracy	Precision	Recall	F1-Score
0.66	0.44	0.66	0.53

AUC Score Average

0.56



6

**Modelo 3.**

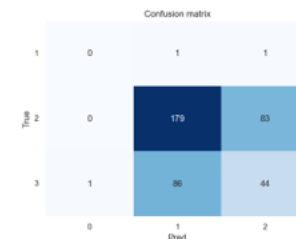
Treino/Teste = 90/10

Árvore de Decisão

Accuracy	Precision	Recall	F1-Score
0.56	0.56	0.56	0.56

AUC Score Average

0.51



7

**Modelo 4.**

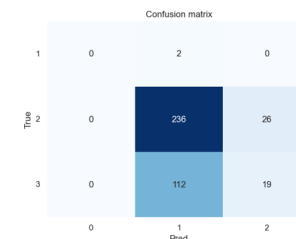
Treino/Teste = 90/10

Random Forest

Accuracy	Precision	Recall	F1-Score
0.65	0.58	0.65	0.57

AUC Score Average

0.60



8