

# Mineração de dados aplicada à dados de evento no futebol

Luis Felipe Ramos   Igor Lacerda Faria da Silva   Matheus  
Tiago Pimenta de Souza

2024-07-28

# Uso de ciência de dados aplicada ao esporte



Moneyball (2011)



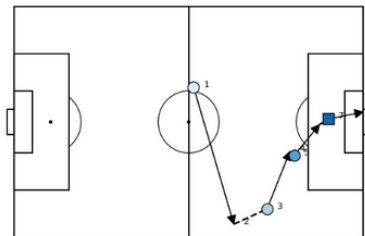
FAME

# Dados de súmula no futebol



Estatísticas da final da Copa do Mundo de 2022

# Dados de evento



	time_seconds	type	player	team
○ 1	1835.0	pass	Rodrigo Hernández Cascante	Manchester City
- - 2	1838.0	dribble	Kyle Walker	Manchester City
○ 3	1840.0	pass	Kyle Walker	Manchester City
- - 4	1841.0	dribble	Kevin De Bruyne	Manchester City
○ 5	1842.0	pass	Kevin De Bruyne	Manchester City
- - 6	1843.0	dribble	Gabriel Fernando de Jesus	Manchester City
■ 7	1845.0	shot	Gabriel Fernando de Jesus	Manchester City

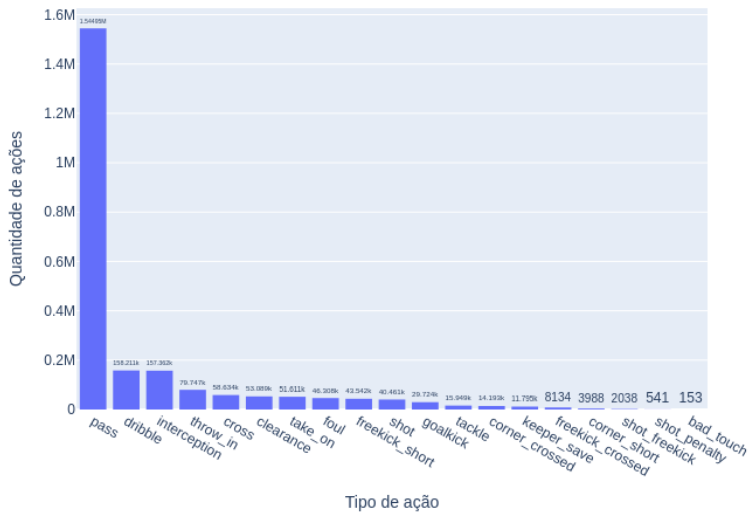
Jogada do Manchester City

# SPADL

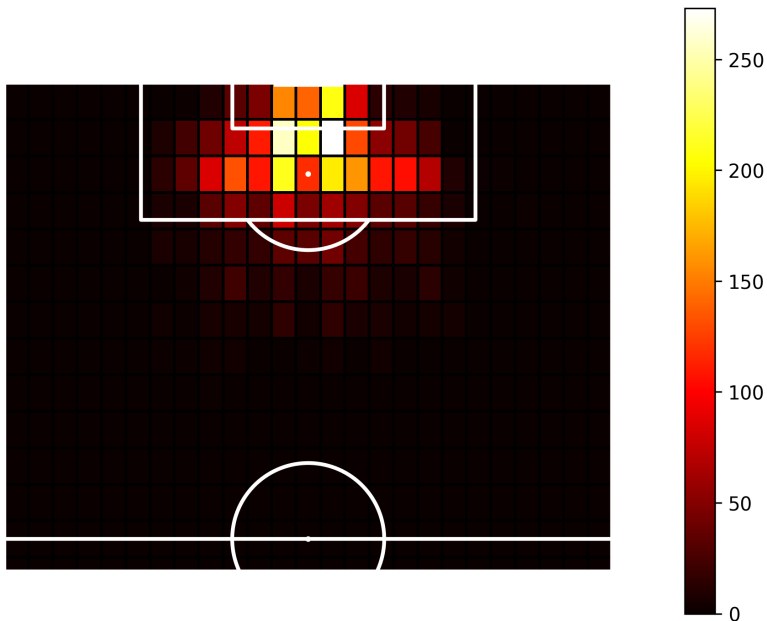
Atributo	Descrição
game_id	O ID do jogo no qual a ação foi realizada
period_id	O ID do período do jogo no qual a ação foi realizada
seconds	O tempo de início da ação
player	O jogador que realizou a ação
team	O time do jogador
start_x	A localização x onde a ação começou
start_y	A localização y onde a ação começou
end_x	A localização x onde a ação terminou
end_y	A localização y onde a ação terminou
action_type	O tipo de ação (por exemplo, passe, chute, drible)
result	O resultado da ação (por exemplo, sucesso ou falha)
bodypart	A parte do corpo do jogador usada para a ação

Table: Descrição dos dados no formato SPADL

# Distribuição dos tipos de ações



## Mapa de calor dos gols



# Objetivos

## Descoberta de subgrupos

- ▶ Trabalho motivador: *Subgroup Discovery in Soccer Data*
- ▶ Reproduzir os experimentos do artigo e experimentar com novos algoritmos de SD



# Objetivos

## Mineração de sequências

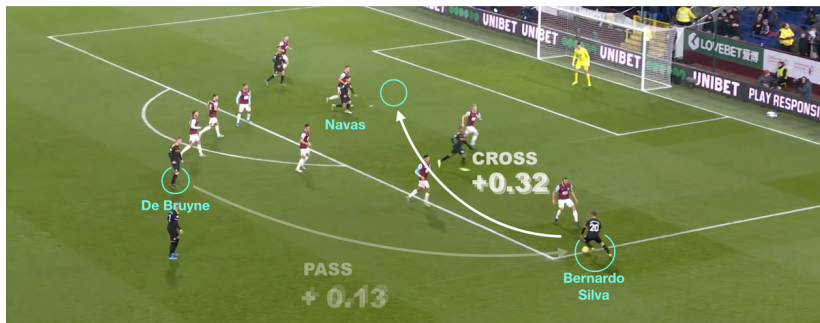
- ▶ Trabalho motivador: *Supervised sequential pattern mining of event sequences in sport to identify important patterns of play: An application to rugby union*
- ▶ Aplicar as ideias em sequências de jogadas de futebol

## Gols Esperados (xG)

- ▶ Métrica de cálculo de probabilidade de gol
- ▶ Diferentes *features* podem ser utilizadas:
  - ▶ Distância para o gol
  - ▶ Ângulo de chute



- ▶ Ações valem mais do que apenas gols
- ▶ O quanto cada jogada contribui para o aumento da chance de gol e diminuição da chance de tomar gol?



# Descoberta de subgrupos

-res

# Mineração de sequências

- ▶ Resultados não muito promissores usando o SPP

# Fim

- ▶ Perguntas, comentários, etc