# Trabalho Final

Prévia

Artur Abreu Hendler - 231713 Guilherme Zeferino Rodrigues Dobins - 236129 João Vitor Baptista Moreira - 237833

#### Retomando:

Contexto gerador e motivação

#### Motivação e Contexto gerador:

Esse projeto tem como objetivo observar a relação entre os odds de apostas de uma partida e seu possível resultado. Atrelando fatores como a popularidade de um jogador, país de confronto e terreno do jogo e estatísticas das apostas do passado com o resultado do confronto.



Tennis Data UK (xls): Resultados de jogos, odds e rankings



Functional Tennis (CSV): Compilado de Tenistas mais populares no Instagram em 2019



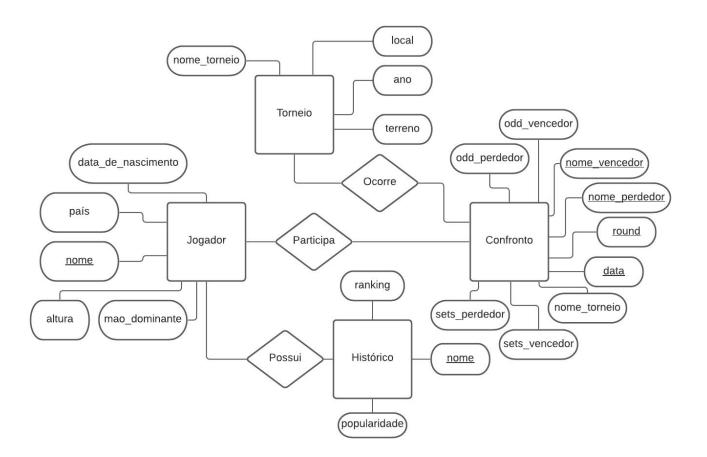
#### Fontes



Instagram (pags. HTML): Número de seguidores atual



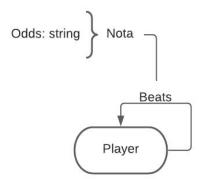
#### Modelo Conceitual

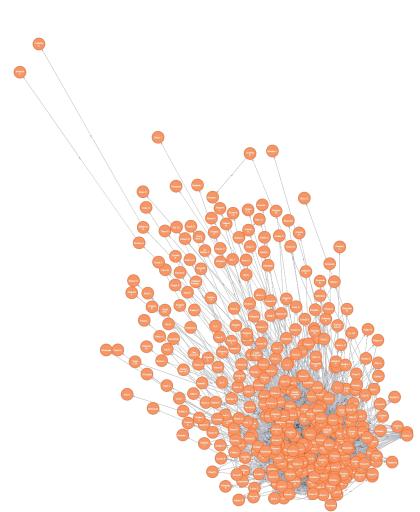


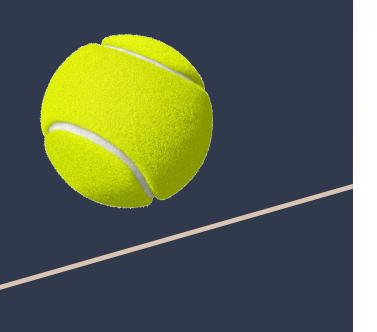


Modelo de Grafo e instância (reduzida):

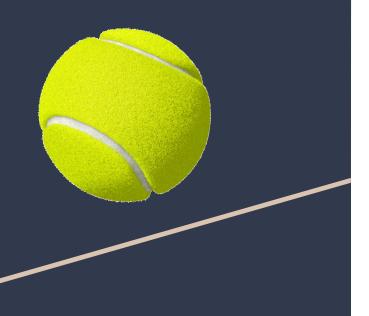
## Grafo



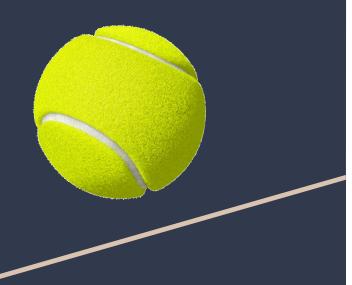




## PERGUNTAS INICIAIS



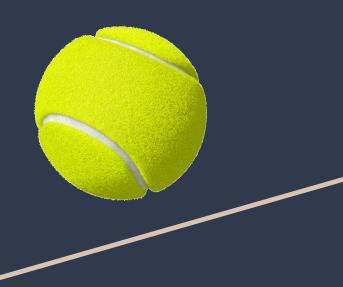
 A altura de um jogador influencia diretamente no seu desempenho nas partidas?



index	NOME	DATA_DE_NASCIMENTO	ALTURA	MAO_DOMINANTE	LOCAL
0	Rochus O.	19810118	168	R	BEL
1	Opelka R.	19970828	211	R	USA

Iindex	WINNER	NUM	
0	Opelka R.	66	
1	Rochus O.	38	

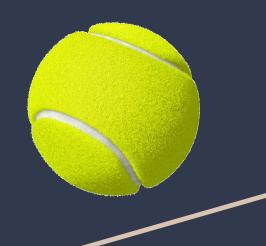
index	LOSER	NUM	
0	Opelka R.	64	
1	Rochus O.	61	



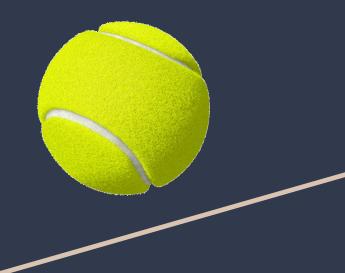
- A altura de um jogador influencia diretamente no seu desempenho nas partidas?
- Em quais torneios os jogadores Rafael Nadal, Novak Djokovic e Roger Federer possuem mais participações?



index	WINNER	TOURNAMENT	NUM
0	Nadal R.	French Open	74
1	Djokovic N.	Australian Open	68
2	Djokovic N.	French Open	65
3	Djokovic N.	Wimbledon	64
4	Djokovic N.	US Open	62
5	Federer R.	Australian Open	56
6	Federer R.	Wimbledon	54
7	Nadal R.	Australian Open	48
8	Djokovic N.	Internazionali BNL d'Italia	48
9	Nadal R.	Internazionali BNL d'Italia	46
10	Nadal R.	US Open	44
11	Nadal R.	Monte Carlo Masters	44
12	Federer R.	Swiss Indoors	43
13	Federer R.	BNP Paribas Open	40
14	Federer R.	US Open	39
15	Djokovic N.	BNP Paribas Open	37
16	Nadal R	Open Banco Sabadell	36

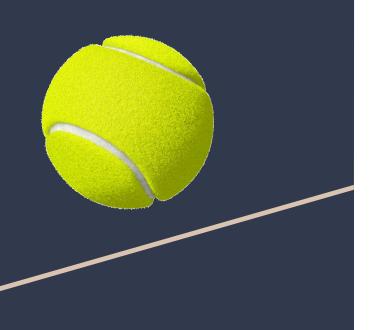


- A altura de um jogador influencia diretamente no seu desempenho nas partidas?
- Em quais torneios os jogadores Rafael Nadal, Novak Djokovic e Roger Federer possuem mais participações?
- Qual a distribuição da mão dominante dos jogadores brasileiros? Quais os nomes dos canhotos?

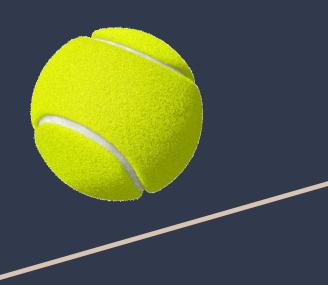


lindex	MAO_DOMINANTE	NUM
0	R	14
1	L	3
2	U	1

index	NOME
0	Mello R.
1	Bellucci T.
2	Monteiro T



## PERGUNTAS FUTURAS



- É possível identificar jogadores mais habilidosos a partir das relações que eles criam ou de suas rivalidades?
- O local do torneio afeta a performance de um jogador?
- Como a idade de um jogador afeta sua chance de avançar no tornejo?

#### Dificuldades encontradas

- Scrapping do instagram bloqueio após algumas tentativas
- Nomes escritos de diferentes formas precisamos transformar em uma pessoa só. Exemplo R. Nadal e Rafael Nadal
- Pessoas com nomes iguais geram conflitos nos grafos
- Dados faltantes na data de nascimento chave primária do jogador
- Ranking masculino e feminino



#### Próximos passos

- Consertar a tabela de confrontos, de modo que não haja nomes diferentes de uma mesma pessoa
- Continuar o scrapping dos dados de popularidade do instagram
- Dividir a base de dados entre torneios e jogadores masculinos e femininos para que o ranking n\u00e3o seja afetado
- Resolver o problema de dados faltantes em algumas colunas (mediana? valor mínimo?)



# Obrigado!