

**UNIVERSIDADE DE BRASÍLIA**  
**Faculdade do Gama**

**Sistemas de Banco de Dados 2**

**Trabalho Final (TF)**

**Mineração de dados (Data Mining)**

**Amanda Jeniffer Pereira Nobre – 190124997**

**Victor Hugo Oliveira Leão – 200028367**

**Vinícius Assumpção de Araújo – 200028472**

Brasília, DF

2022

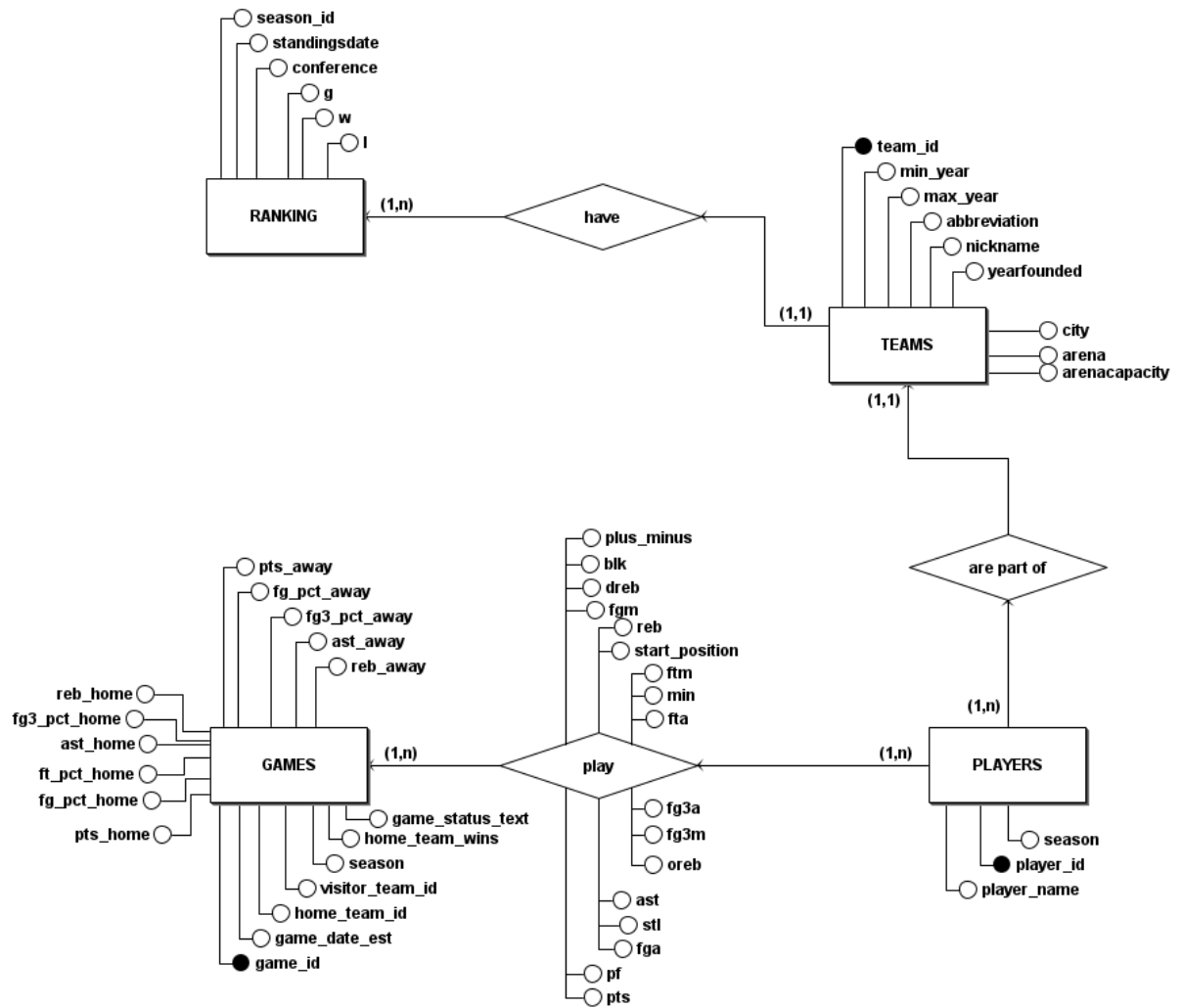
## Contextualização e Base de Dados

A tecnologia tema do grupo é Data Mining, que consiste em um processo de exploração em grandes quantidades de dados tentando buscar padrões, conexões, relacionamentos e/ou regras entre diferentes variáveis e atributos. Assim, em outras palavras, podemos dizer que o Data Mining transforma os dados em conhecimento e, portanto, é uma técnica pouco engessada, podendo ser usada com sucesso em diversos contextos.

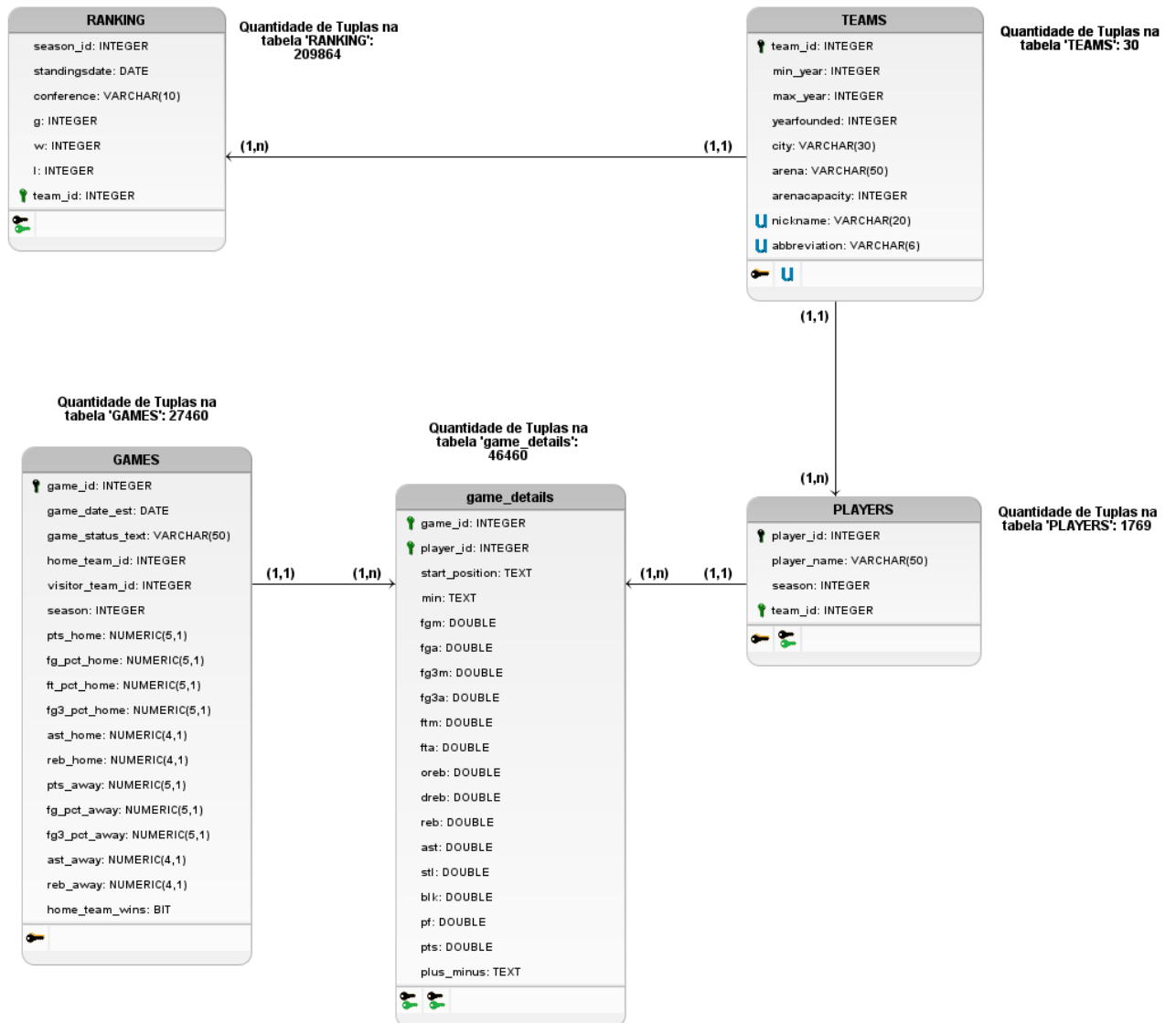
Nesse sentido, a base de dados escolhida foi a da **NBA Games** (National Basketball Association), a liga nacional americana de basquetebol, sugerida por **Victor Hugo Oliveira Leão**. Essa base de dados possui dados de times, seus rankings ao longo do tempo, jogadores desses times e jogos em que esses jogadores participaram, com suas estatísticas em cada jogo. Com isso, o Data Mining pode ser aplicado buscando encontrar diversos padrões e extrair conhecimento de muitas maneiras nessa base de dados, já que ela possui 5 tabelas ligadas por relacionamentos, sendo concisa.

A base de dados pode ser encontrada no seguinte local virtual:  
<https://www.kaggle.com/datasets/nathanlauga/nba-games>.

# DE-R



# DLD



# DICIONÁRIO DE DADOS

Entidade: TEAMS				
Descrição: Define os times da liga.				
Atributo	Propriedades do atributo	Tipo de dado	Tamanho	Descrição
team_id	chave primária obrigatório	integer	-	Identificador do time
min_year	obrigatório	integer	-	Primeiro ano que o time jogou na liga
max_year	obrigatório	integer	-	Último ano que o time jogou na liga
abbreviation	obrigatório único	varchar	6	Abreviação do nome do time
nickname	obrigatório único	varchar	20	Apelido do time (normalmente segundo nome)
yearfounded	obrigatório	integer	-	Ano de fundação do time
city	obrigatório	varchar	30	Cidade natal do time
arena	obrigatório	varchar	50	Nome do ginásio do time
arenacapacity	obrigatório	integer	-	Capacidade de espectadores do ginásio

Entidade: RANKING				
Descrição: Define o ranking de um time em um determinado dia.				
Atributo	Propriedades do atributo	Tipo de dado	Tamanho	Descrição
season_id	obrigatório	integer	-	Identificador da temporada do Ranking
standingsdate	obrigatório	date	-	Data do ranking
conference	obrigatório	varchar	10	Identifica qual conferência o time faz parte

g	obrigatório	integer	-	Número de jogos jogados na temporada
w	obrigatório	integer	-	Número de jogos ganhos na temporada
l	obrigatório	integer	-	Número de jogos perdidos na temporada
team_id	obrigatório chave estrangeira	integer	-	Identificador do time do ranking (chave primária em TEAMS)

Entidade: PLAYERS				
Descrição: Define os jogadores de um time.				
Atributo	Propriedades do atributo	Tipo de dado	Tamanho	Descrição
player_id	chave primária obrigatório	integer	-	Identificador do jogador
player_name	obrigatório	varchar	50	Nome do jogador
season	obrigatório	integer	-	Temporada que o jogador jogou
team_id	obrigatório chave estrangeira	integer	-	Identificador do time que o jogador joga (chave primária em TEAMS)

Entidade: GAMES				
Descrição: Define os jogos que ocorreram na liga				
Atributo	Propriedades do atributo	Tipo de dado	Tamanho	Descrição
game_id	obrigatório chave primária	integer	-	Identificador do jogo
game_date_est	obrigatório	date	-	Data do jogo
game_status_text	obrigatório	varchar	50	Status do jogo (finalizado ou não)
home_team_id	obrigatório	integer	-	Identificador do time da casa
visitor_team_id	obrigatório	integer	-	Identificador do time visitante

season	obrigatório	integer	-	Temporada em que o jogo ocorreu
pts_home	-	numeric	5,1	Quantidade de pontos que o time da casa fez
fg_pct_home	-	numeric	5,1	Porcentagem de arremessos acertados pelo time da casa
ft_pct_home	-	numeric	5,1	Porcentagem de lances livres acertados pelo time da casa
fg3_pct_home	-	numeric	5,1	Porcentagem de arremessos de 3 pontos acertados pelo time da casa
ast_home	-	numeric	4,1	Quantidade de assistências feitas pelo time da casa
reb_home	-	numeric	4,1	Quantidade de rebotes feitos pelo time da casa
pts_away	-	numeric	5,1	Quantidade de pontos que o time visitante fez
fg_pct_away	-	numeric	5,1	Porcentagem de arremessos acertados pelo time visitante
fg3_pct_away	-	numeric	5,1	Porcentagem de arremessos de 3 pontos acertados pelo time visitante
ast_away	-	numeric	4,1	Quantidade de assistências feitas pelo time visitante
reb_away	-	numeric	4,1	Quantidade de rebotes feitos

				pelo time visitante
home_teams_wins	obrigatório	numeric	bit	Indica se o time da casa ganhou (1) ou não (0) a partida

Relacionamento: game_details				
<b>Descrição:</b> Define relacionamento que corresponde aos jogadores que jogaram um jogo com cardinalidade n:m.				
Atributo	Propriedades do atributo	Tipo de dado	Tamanho	Descrição
player_id	chave estrangeira obrigatório	integer	-	Identificador do jogador que jogou um jogo (chave primária em PLAYERS)
game_id	chave estrangeira obrigatório	integer	-	Identificador do jogo (chave primária em GAMES)
start_position	nulo como padrão	text	-	Posição do jogador titular
min	nulo como padrão	text	-	Minutos que o jogador jogou
fgm	nulo como padrão	double	-	Quantidade de cestas acertadas
fga	nulo como padrão	double	-	Quantidade de cestas tentadas
fg3m	nulo como padrão	double	-	Quantidade de cestas de 3 pontos acertadas
fg3a	nulo como padrão	double	-	Quantidade de cestas de 3 pontos tentadas
ftm	nulo como padrão	double	-	Quantidade de lances livres acertados



fta	nulo como padrão	double	-	Quantidade de lances livres tentados
oreb	nulo como padrão	double	-	Quantidade de rebotes ofensivos
dreb	nulo como padrão	double	-	Quantidade de rebotes defensivos
reb	nulo como padrão	double	-	Quantidade total de rebotes
ast	nulo como padrão	double	-	Quantidade de assistências
stl	nulo como padrão	double	-	Quantidade de roubo de bolas
blk	nulo como padrão	double	-	Quantidade de bloqueios
pf	nulo como padrão	double	-	Quantidade de faltas pessoais
pts	nulo como padrão	double	-	Quantidade de pontos
plus_minus	nulo como padrão	text	-	Plus Minus (estatística relacionada ao time quando o jogador está em quadra)

View: RANKING_SEASON2021		
<b>Descrição:</b> View que apresenta os rankings na temporada 2021-2022 da NBA para todos os times.		
Atributo	Tabela Base	Descrição
nickname	TEAMS	Apelido do time
standingsdate	RANKING	Data em que o ranking/classificação (números de jogos ganhos e perdidos) está sendo apresentada
g	RANKING	Número de jogos jogados na data do ranking
w	RANKING	Número de jogos ganhos na data do ranking
l	RANKING	Número de jogos perdidos na data do ranking