

Análise de Grafos no Azure SQL DB

Wagner Crivelini

Patrocínio:

GOLD



SILVER



Apoiadores:



BRONZE



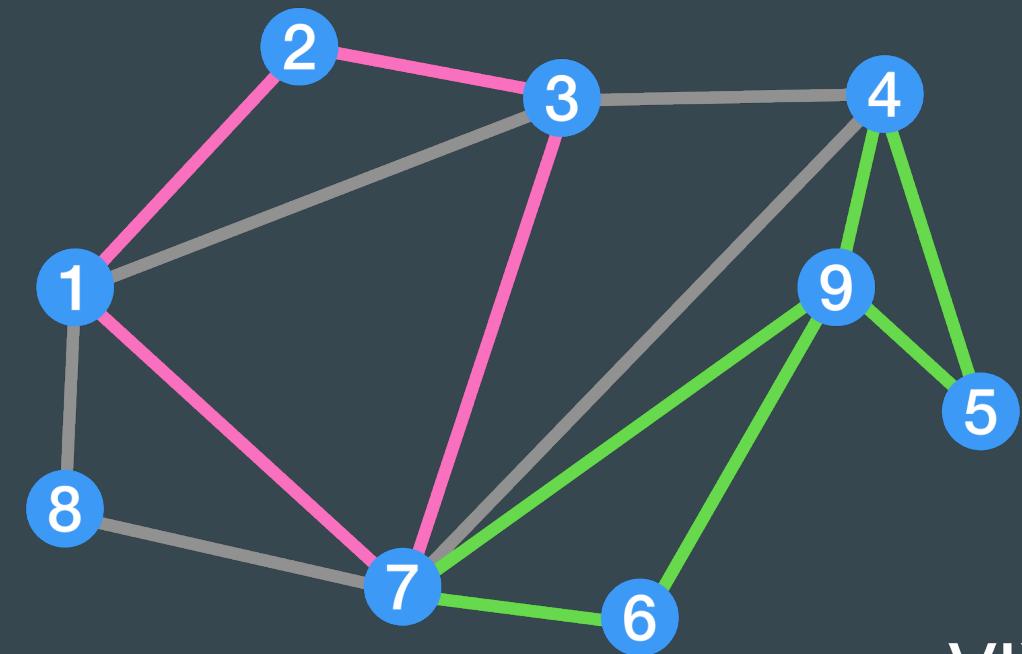
BIO: Wagner Crivelini

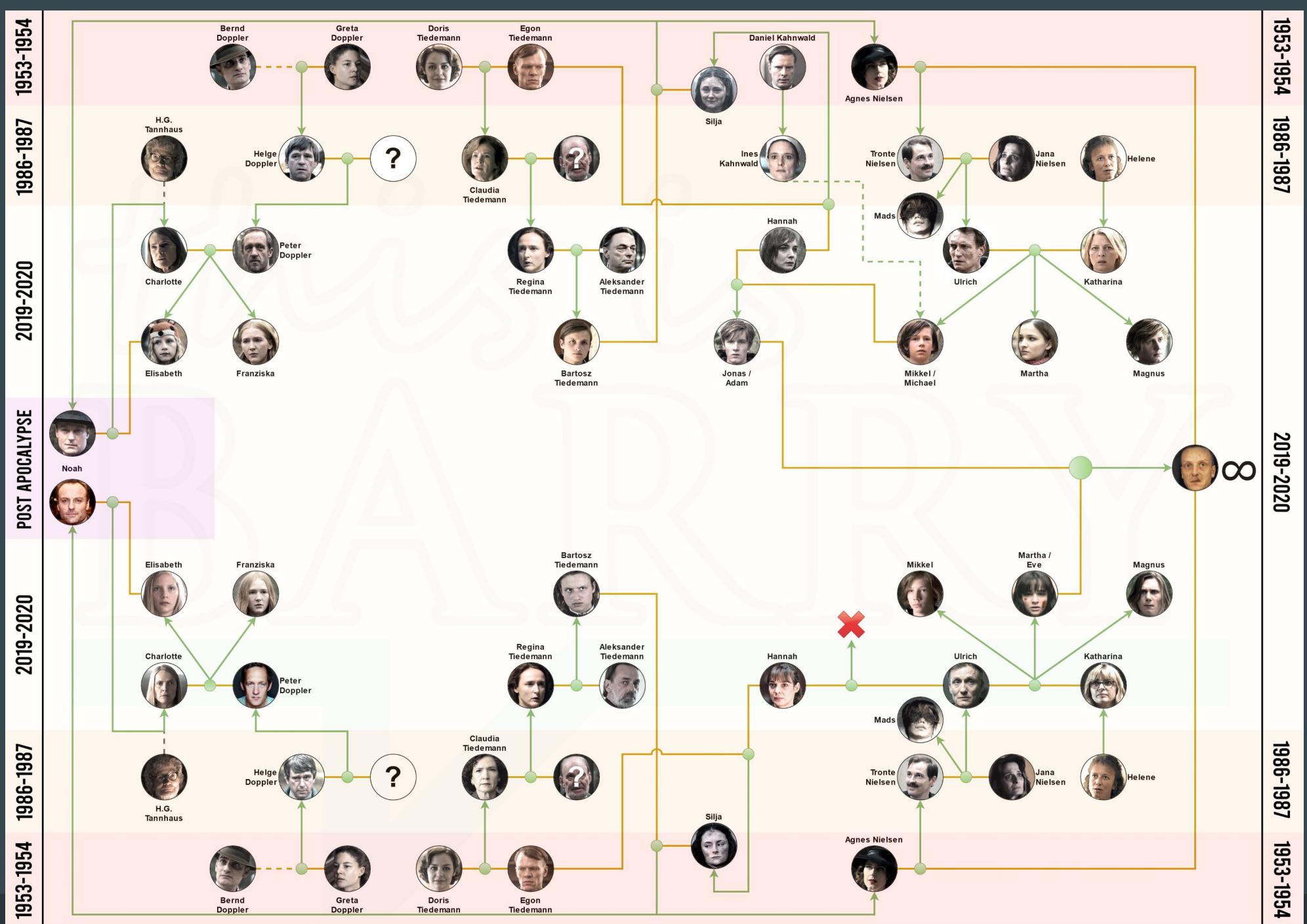
- Consultor de Data & AI da MICROSOFT ISD BRASIL
- Colunista em diversos portais (SQLServerCentral, iMasters)
 - +150 artigos
 - +100 podcasts
 - +050 palestras



O Que São Grafos

- Modelos que lidam com **Relações** complexas entre **Entidades**
- Usados em todo tipo de negócio: social media, engenharia...
- *"6 Graus de Separação"*
- Exemplo:
 - LINKEDIN: classificação de contatos

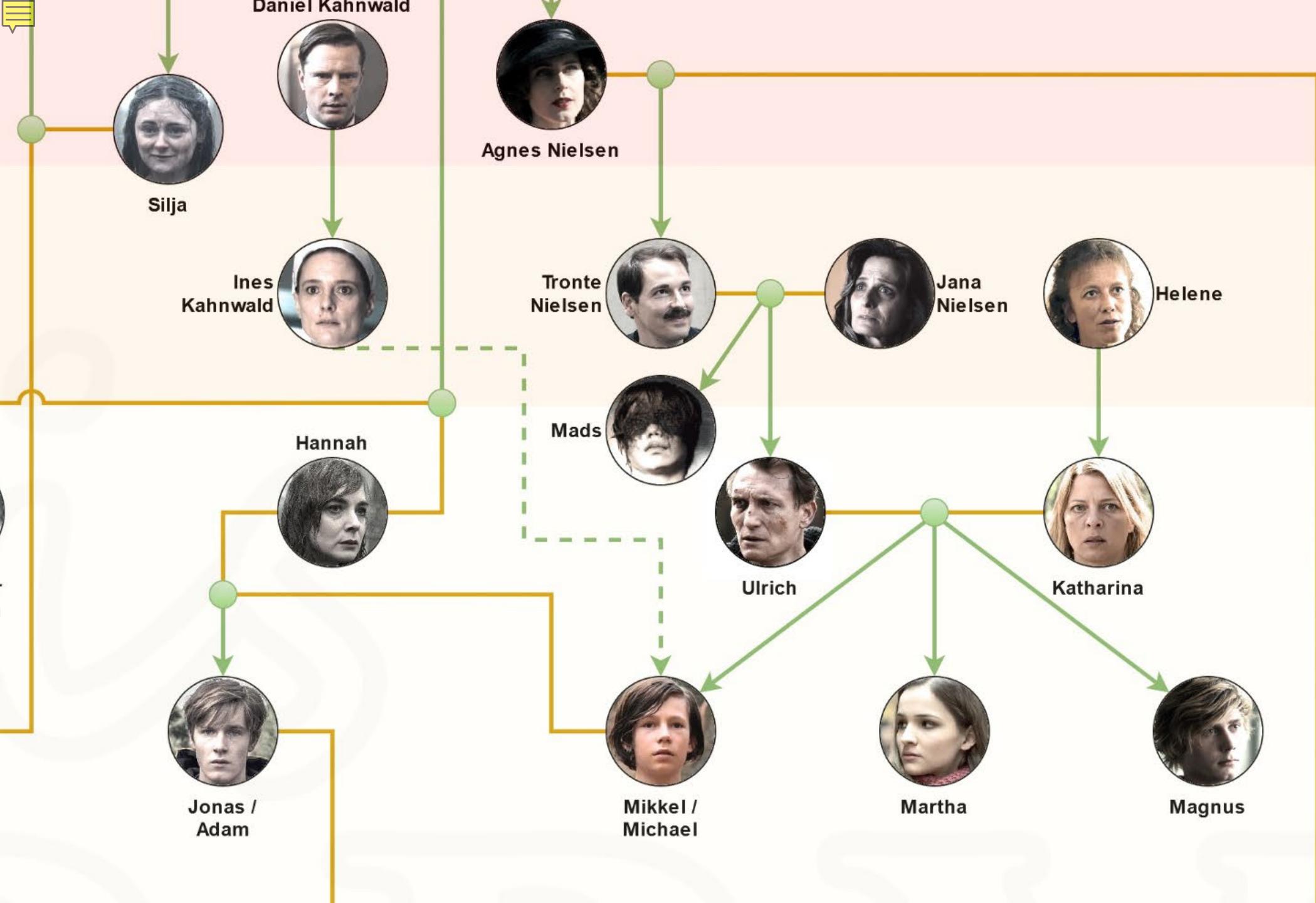




VIX25

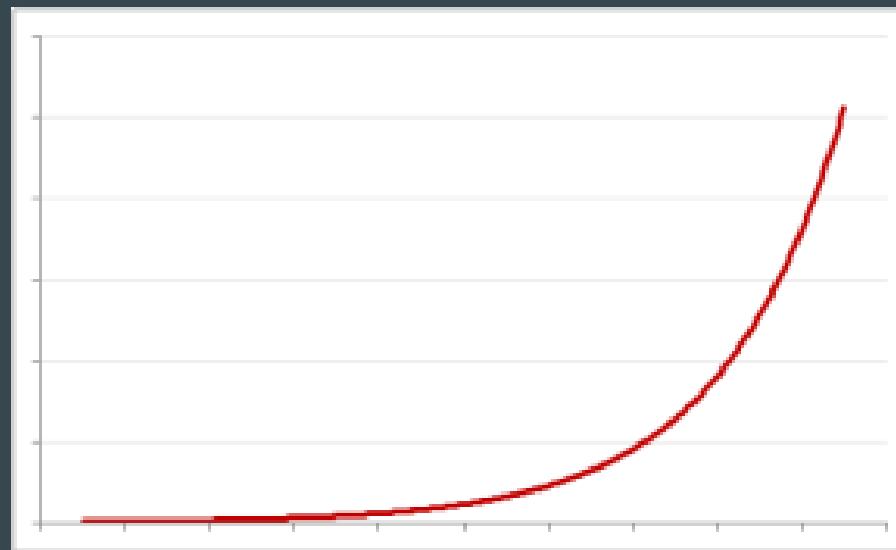
www.dataes.com.br

53-1954 1986-1987



Onde Aplicar Modelos de Grafos

- Modelos com relacionamentos complexos
- Quando ENTIDADES se apresentam com grande número de elementos distintos





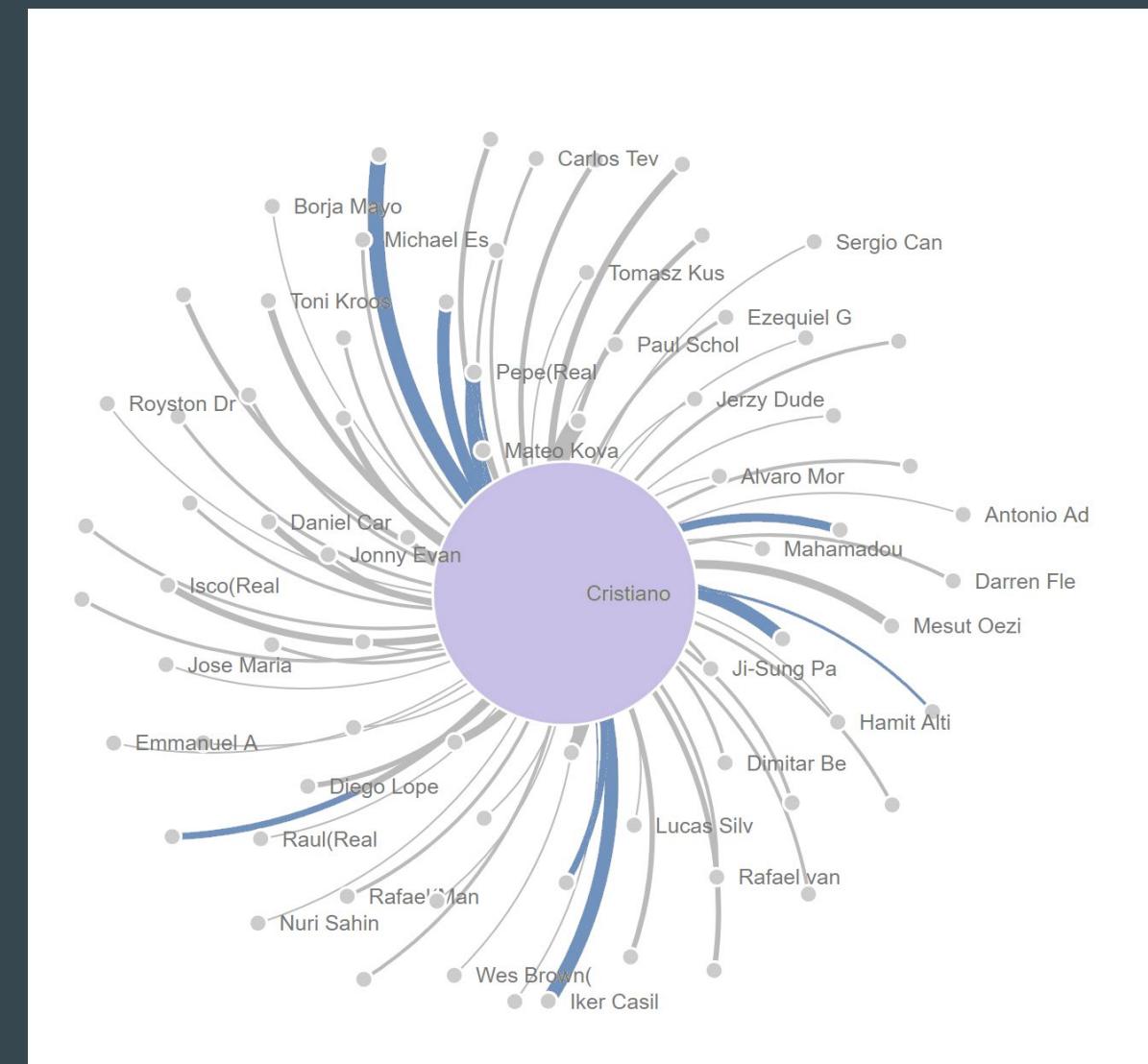
BD de Grafos = BD Relacional?

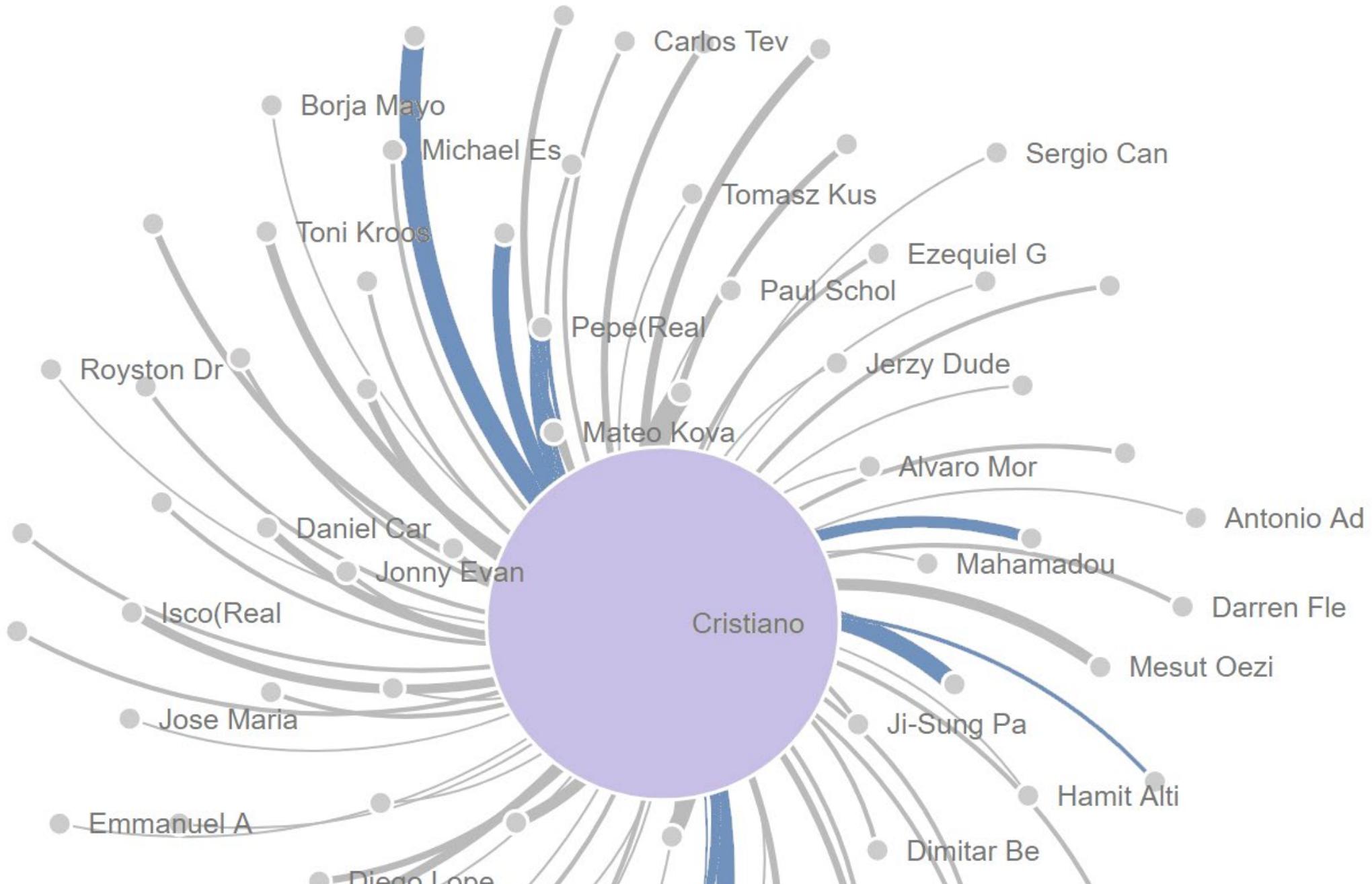
- São bastante diferentes
- **Entidades** no modelo relacional são tabelas e nos grafos cada elemento da tabela é uma entidade.
- Grafos tratam de **Relações**, mas estas podem ser mais complexas e numerosas que as definidas no modelo relacional
- Relacional: raramente passa de 20 relações entre tabelas
- Grafos: uma entidade pode facilmente ter milhares relações



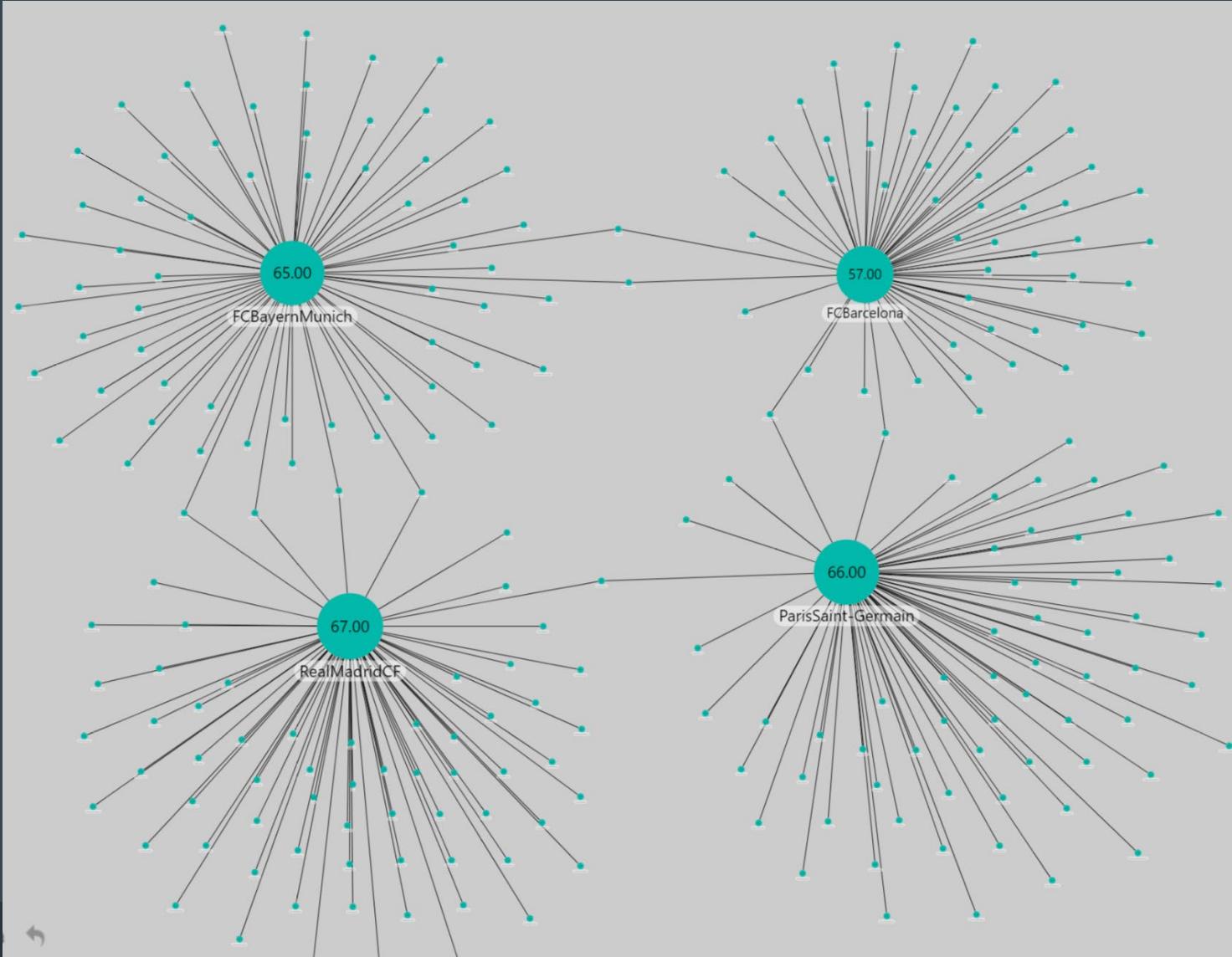
Um Único Grafo

- Número de jogos que cada jogador fez com ou contra Cristiano Ronaldo
- Linhas grossas representam maior número de jogos

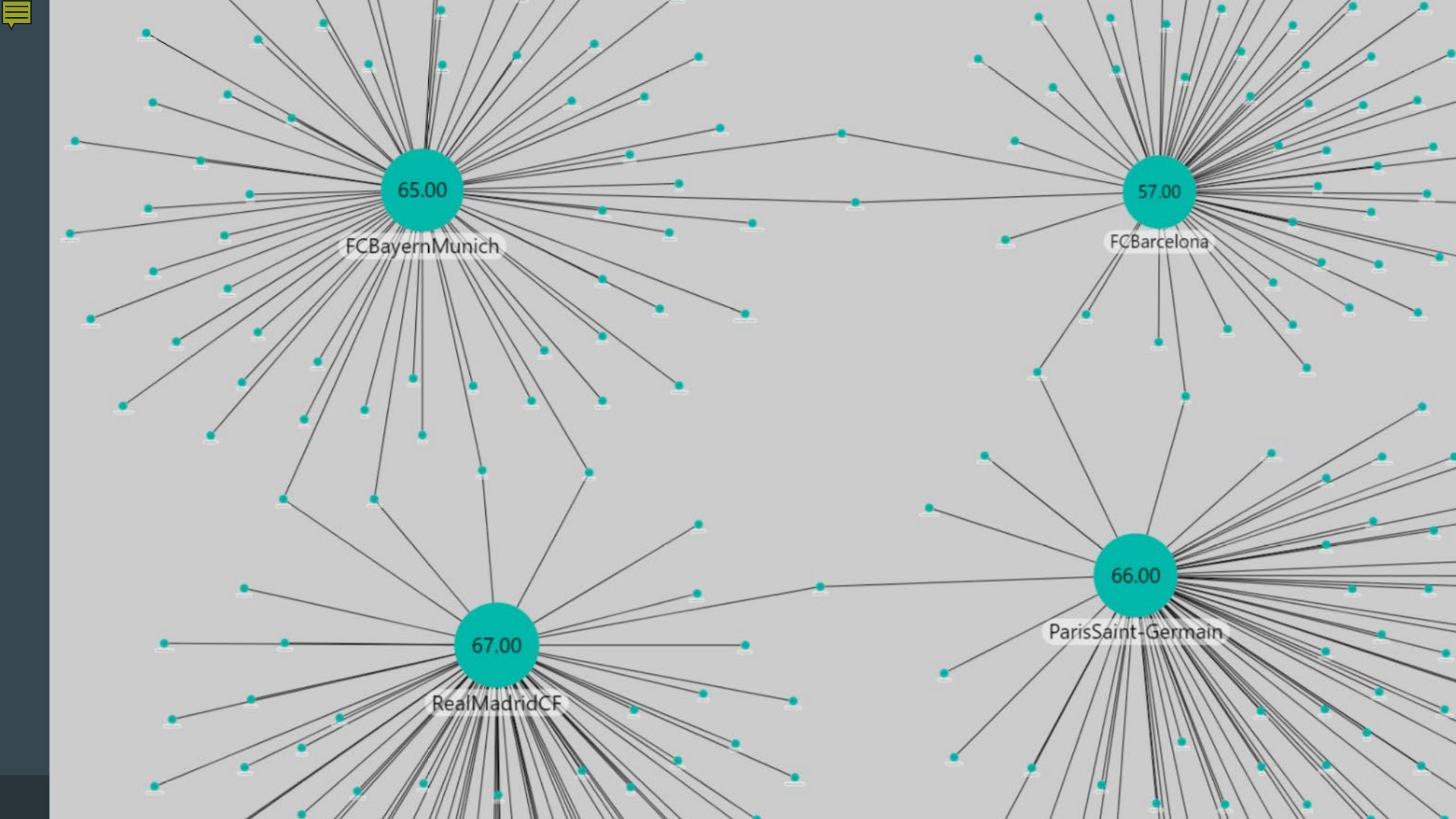




Relacionamentos entre Grafos



Jogadores que
atuaram por cada
Time num período de
N temporadas



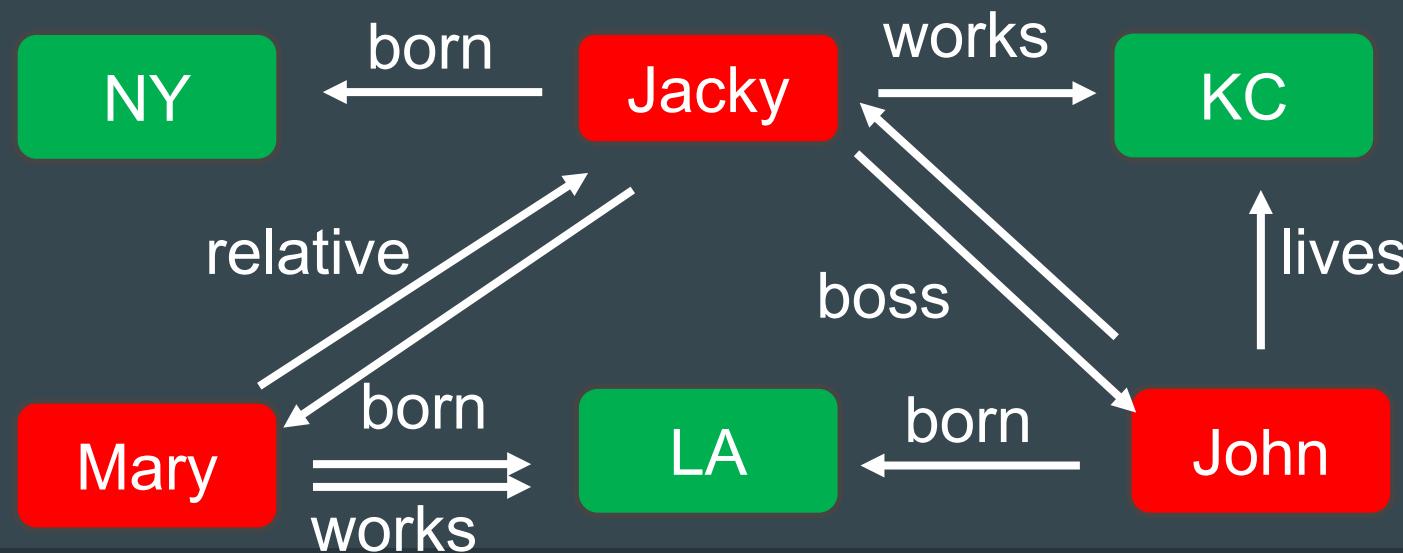
Definição do Modelo

Grafos e T-SQL

- Disponíveis desde 2017 (On-Premises e na nuvem)
- Visão lógica de dados
 - Tabelas de nós (NODE) lidam com ENTIDADES
 - Tabelas de bordas (EDGE) tratam de RELACIONAMENTOS
- Porta de entrada para universo dos grafos
- Carga de dados para modelo de grafos requer cuidado

Exemplo de Modelo de Grafos

- ENTIDADES: Cidades (verde), Pessoas (Vermelho)
- RELACIONAMENTOS: onde nasceu, onde mora, onde trabalha, quem é parente, quem é chefe



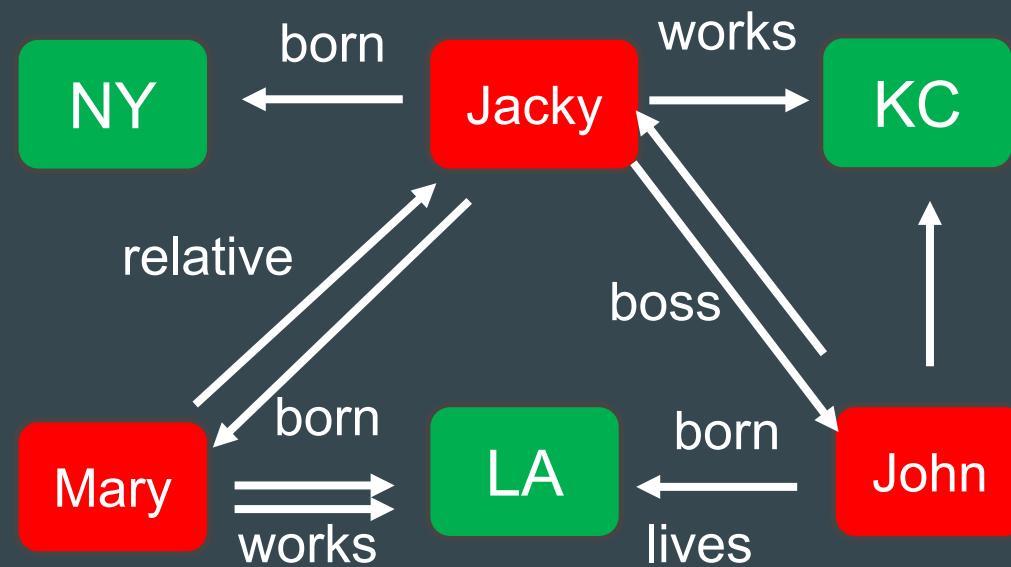
Definição do Modelo de Grafos

1. NODES (i.e. Entities)

- a) People (Jacky, John, Mary)
- b) Cities (NY, LA, KC)

2. EDGES (i.e. Relationships)

- a) Born in
- b) Works in
- c) Lives in
- d) Relative
- e) Boss





Implementação do Modelo de Grafos

- Importante:
 - **BORDAS** são definidas com uma **DIREÇÃO**

Person

WAS BORN IN

City

Person

WORKS IN

City

Person1

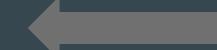
IS A RELATIVE

Person2

Person2

IS A RELATIVE

Person1



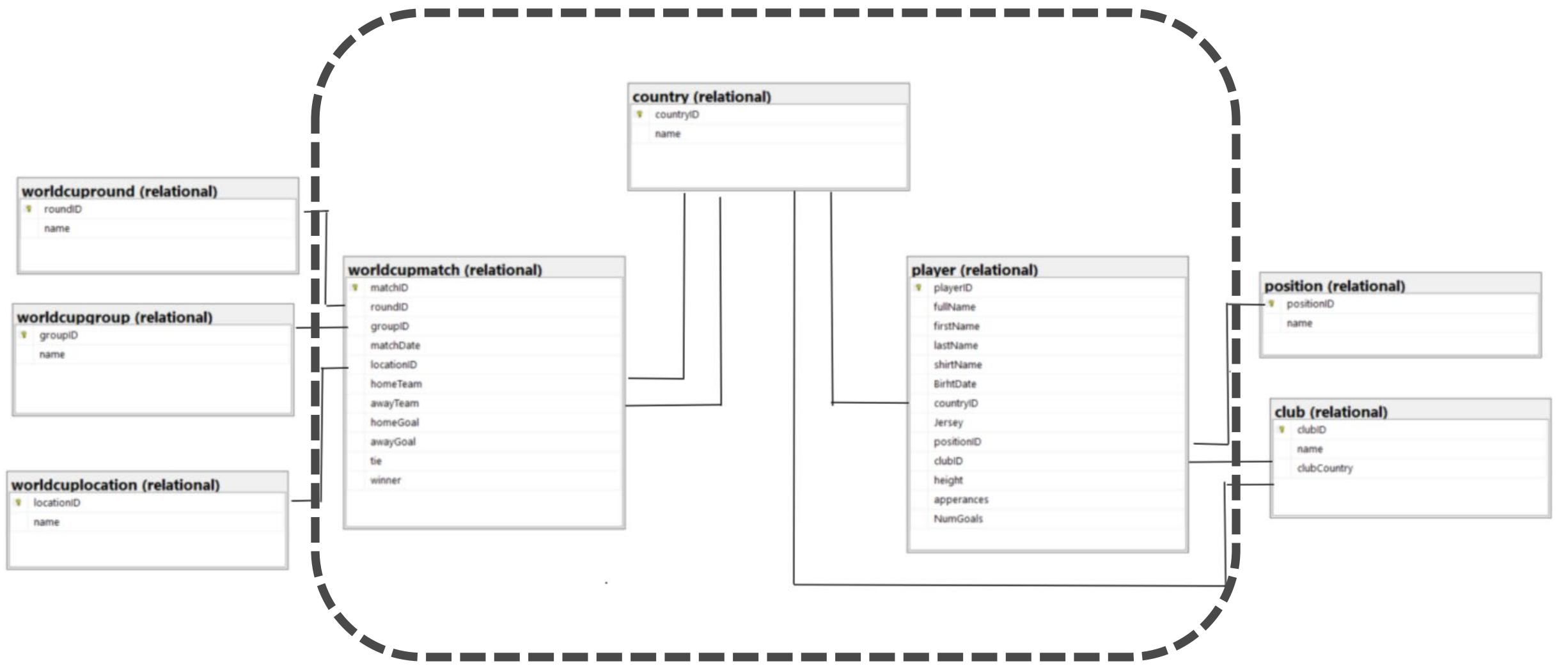
Rel. Bidirecionais:
tratar no **modelo** ou
então na **consulta**

Mão na Massa

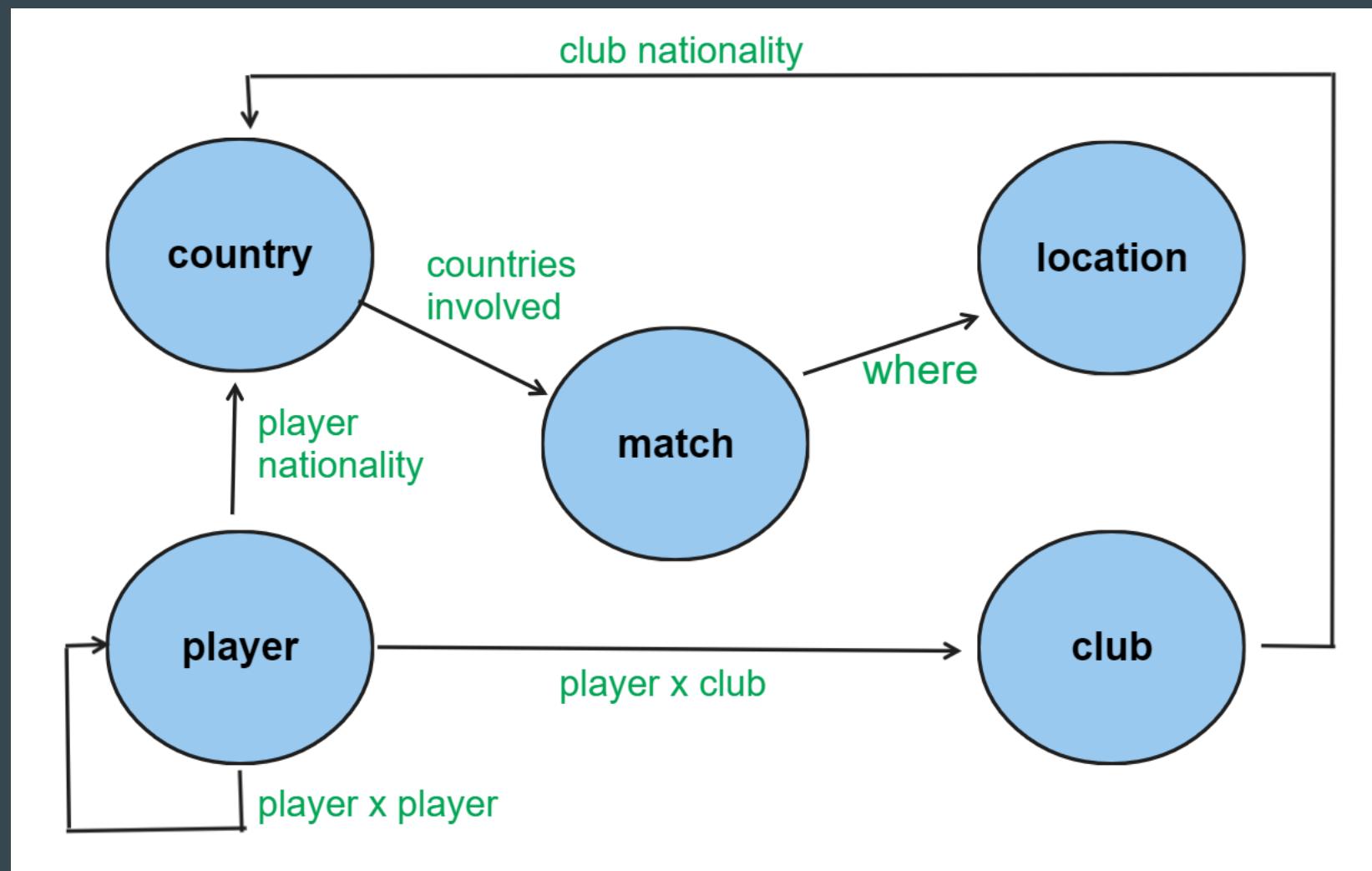
- Dataset: **Copa do Mundo 2022** (FIFA)
- 32 equipes
- 831 atletas
- 296 clubes



FIFA 2022 – Modelo Relacional



FIFA 2022 – Modelo de Grafos



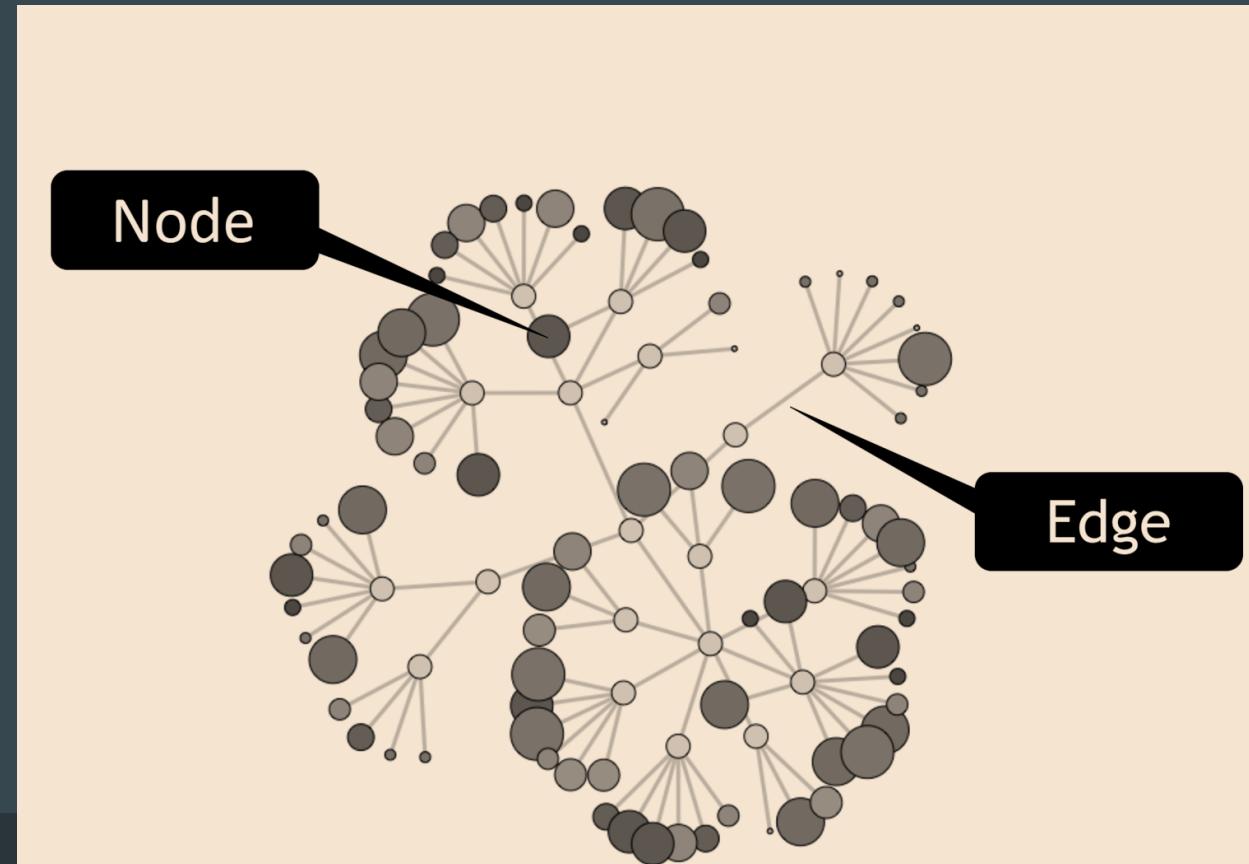
Implementação do Modelo

VIX25

www.dataes.com.br 

Grafos e T/SQL

- Criando tabelas de nós e bordas no Azure SQL DB/SQL Server 2025
- Carga de dados
- Consultas





DDL para Grafos

- Tabelas de **Nós** usam CREATE TABLE (...) AS NODE
- **Bordas** usam CREATE TABLE [(.)] AS EDGE
- Os nós criam automaticamente uma coluna \$node_id

```
{"type":"node","schema":"node","table":"club","id":0}
```



Demo

VIX25

Carga de Dados para Tabelas de Grafos

- INSERT normal para tabelas de **NÓS**
- INSERT nas **BORDAS** deve trazer os **\$node_id** das tabelas envolvidas mais os atributos adicionais

```
INSERT INTO edge.clubXcountry
SELECT NCL.$node_id , NCO.$node_id
FROM node.club NCL
INNER JOIN relational.club RC ON NCL.ID = RC.ID
INNER JOIN node.country NCO ON CO.ID = RC.ID
GO
```



Demo

VIX25

Análises com Grafos

VIX25

www.dataes.com.br 

Consultando Dados de Grafos

- Principal diferença:
Não existe JOIN.

- O que se usa:
Operador **Match** na cláusula **WHERE**

Ex:

```
WHERE MATCH (t1-(t2)->t3)
```



Demo

VIX25

Observe a DIREÇÃO

- Operador MATCH usa ao menos 1 nó e 1 borda
- Mas o MATCH define a direção da busca de dados

```
SELECT .....  
FROM nodes.Player P1, edges.PlayersInvolved PIN, nodes.Game M  
WHERE MATCH (t1-(t2)->t3)
```

É idêntico a

```
SELECT .....  
FROM nodes.Player P1, edges.PlayersInvolved PIN, nodes.Game M  
WHERE MATCH (t3 <-(t2)-t1)
```



DEMO 1: Consultas com Grafos

- Exemplo: Lista de jogos com times mandantes (HOME)

```
SELECT M.matchID, m.matchDate, m.[group], m.round  
, score = C1.countryID + ' ' + CAST( M.homeGoal AS VARCHAR(2)) + ' x '  
+ CAST( M.awayGoal AS VARCHAR(2)) + '"visitante"'  
, resultado = CASE WHEN M.homeGoal > M.awayGoal THEN 'GANHOU' ELSE  
CASE WHEN M.homeGoal < M.awayGoal THEN 'PERDEU' ELSE 'EMPATOU' END  
END  
FROM node.country C1  
, edge.matchXcountry MC1  
, node.worldcupmatch M  
WHERE MATCH (M - (MC1)-> C1)  
AND MC1.teamStatus = 'home'
```

matchID	matchDate	group	round	score	resultado
1	2022-11-20 16:00:00.000	Group A	1	QAT 0 x 2 "visitante"	PERDEU
2	2022-11-21 16:00:00.000	Group A	1	SEN 0 x 2 "visitante"	PERDEU
3	2022-11-21 13:00:00.000	Group B	1	ENG 6 x 2 "visitante"	GANHOU
4	2022-11-21 19:00:00.000	Group B	1	USA 1 x 1 "visitante"	EMPATOU
5	2022-11-22 19:00:00.000	Group D	1	FRA 4 x 1 "visitante"	GANHOU
6	2022-11-22 13:00:00.000	Group D	1	DEN 0 x 0 "visitante"	EMPATOU
7	2022-11-22 16:00:00.000	Group C	1	MEX 0 x 0 "visitante"	EMPATOU
8	2022-11-22 10:00:00.000	Group C	1	ARG 1 x 2 "visitante"	PERDEU
9	2022-11-23 19:00:00.000	Group F	1	BEL 1 x 0 "visitante"	GANHOU
10	2022-11-23 16:00:00.000	Group E	1	ESP 7 x 0 "visitante"	GANHOU



DEMO 2: Consultas com Grafos

- Exemplo: Jogos com times mandantes e visitantes

```
SELECT M.matchID, m.matchDate, m.[group], m.round  
, score = C1.countryID + ' ' + CAST( M.homeGoal AS VARCHAR(2)) + ' ' + C2.countryID  
+ CAST( M.awayGoal AS VARCHAR(2)) + "visitante"  
, resultado = CASE WHEN M.homeGoal > M.awayGoal THEN 'GANHOU' ELSE  
CASE WHEN M.homeGoal < M.awayGoal THEN 'PERDEU' ELSE 'EMPATOU' END  
END  
FROM node.country C1  
, edge.matchXcountry MC1  
, node.worldcupmatch M  
, node.country C2  
, edge.matchXcountry MC2  
WHERE MATCH (C1 <- (MC1)- M - (MC2) -> C2)  
AND MC1.teamStatus = 'home'  
AND MC2.teamStatus = 'away'
```

matchID	matchDate	group	round	score	resultado
1	2022-11-20 16:00:00.000	Group A	1	QAT 0 x 2 ECU	PERDEU
2	2022-11-21 16:00:00.000	Group A	1	SEN 0 x 2 NED	PERDEU
3	2022-11-21 13:00:00.000	Group B	1	ENG 6 x 2 IRN	GANHOU
4	2022-11-21 19:00:00.000	Group B	1	USA 1 x 1 WAL	EMPATOU
5	2022-11-22 19:00:00.000	Group D	1	FRA 4 x 1 AUS	GANHOU
6	2022-11-22 13:00:00.000	Group D	1	DEN 0 x 0 TUN	EMPATOU
7	2022-11-22 16:00:00.000	Group C	1	MEX 0 x 0 POL	EMPATOU
8	2022-11-22 10:00:00.000	Group C	1	ARG 1 x 2 KSA	PERDEU
9	2022-11-23 19:00:00.000	Group F	1	BEL 1 x 0 CAN	GANHOU
10	2022-11-23 16:00:00.000	Group E	1	ESP 7 x 0 CRC	GANHOU

Relacionamentos Duplos

- Exemplo: COUNTRY vs MATCH
- Partida: Mandante vs Visitante



```
• WHERE MATCH (C1<-(MC1)-M-(MC2)-> C2)  
• AND MC1.teamStatus = 'home'  
• AND MC2.teamStatus = 'away'
```

AND M.SEASON = '2009/2010'



Demo

VIX25

Um Caso Mais Complicado

- Classificar relacionamento entre 1 jogador e demais atletas
- Relação de 1º Grau: Jogadores que estão na Copa E são companheiros de **seleção ou clube** do jogador avaliadado
- Relação de 2º Grau: Jogadores que estão na Copa E jogam em **seleção ou clube** com algum atleta do nível 1
- E assim por diante.



Demo

VIX25

O Caso Complicado – em SQL puro

- Gerou uma consulta com cerca de 90 linhas para apenas dois níveis de relacionamento

Player0	GRAU	Player	PathTo
MESSI	1	RULLI Geronimo (ARG)	MESSI Lionel (ARG)==>RULLI Geronimo (ARG)
MESSI	1	SARABIA Pablo (ESP)	MESSI Lionel (ARG)==>SARABIA Pablo (ESP)
MESSI	1	SOLER Carlos (ESP)	MESSI Lionel (ARG)==>SOLER Carlos (ESP)
MESSI	1	TAGLIAFICO Nicolas (ARG)	MESSI Lionel (ARG)==>TAGLIAFICO Nicolas (ARG)
MESSI	1	VITINHA (POR)	MESSI Lionel (ARG)==>VITINHA (POR)
MESSI	2	ABOUKHLAL Zakaria (MAR)	MESSI Lionel (ARG)==>HAKIMI Achraf (MAR)==>ABOUKHLAL Zakaria (MAR)
MESSI	2	AGUERD Nayef (MAR)	MESSI Lionel (ARG)==>HAKIMI Achraf (MAR)==>AGUERD Nayef (MAR)
MESSI	2	AGUILERA Brandon (CRC)	MESSI Lionel (ARG)==>NAVAS Keylor (CRC)==>AGUILERA Brandon (CRC)
MESSI	2	AKANJI Manuel (SUI)	MESSI Lionel (ARG)==>ALVAREZ Julian (ARG)==>AKANJI Manuel (SUI)
MESSI	2	AKE Nathan (NED)	MESSI Lionel (ARG)==>ALVAREZ Julian (ARG)==>AKE Nathan (NED)

Demo

VIX25



Função SHORTEST_PATH

- Quando não se sabe a “distância” do relacionamento
- “Amigo de um amigo de um amigo”
- Relacional é excessivamente complexo nesses casos

```
WHERE MATCH(SHORTEST_PATH(no1(-(borda)->no2)+))  
AND no1.Player = 'Player1'
```



http://bit.ly/sql_shortest_path

VIX25

O Caso Complicado – em SQL + Grafos

- Gerou uma consulta com 10 linhas SEM LIMITE de níveis de relacionamento

```
-- mesma consulta usando GRAFOS
DECLARE @player VARCHAR(200) = 'MESSI'
SELECT      Player          = NP1.player
            , GRAU           = COUNT(NP2.Player)
            , Player2        = LAST_VALUE(NP2.Player)
            , PathTo         = NP1.player + ' ==> ' + STRING_AGG(NP2.Player, ' ==> ') WITHIN GROUP (GRAPH PATH)
            ,              WITHIN GROUP (GRAPH PATH)
            ,              WITHIN GROUP (GRAPH PATH)
FROM        node.player NP1
            , edge.playerXplayer FOR PATH AS EPP
            , node.player      FOR PATH AS NP2
WHERE       MATCH(SHORTEST_PATH(NP1(-(EPP)->NP2)+))
            AND NP1.shirtName = @player
```

Demo

VIX25

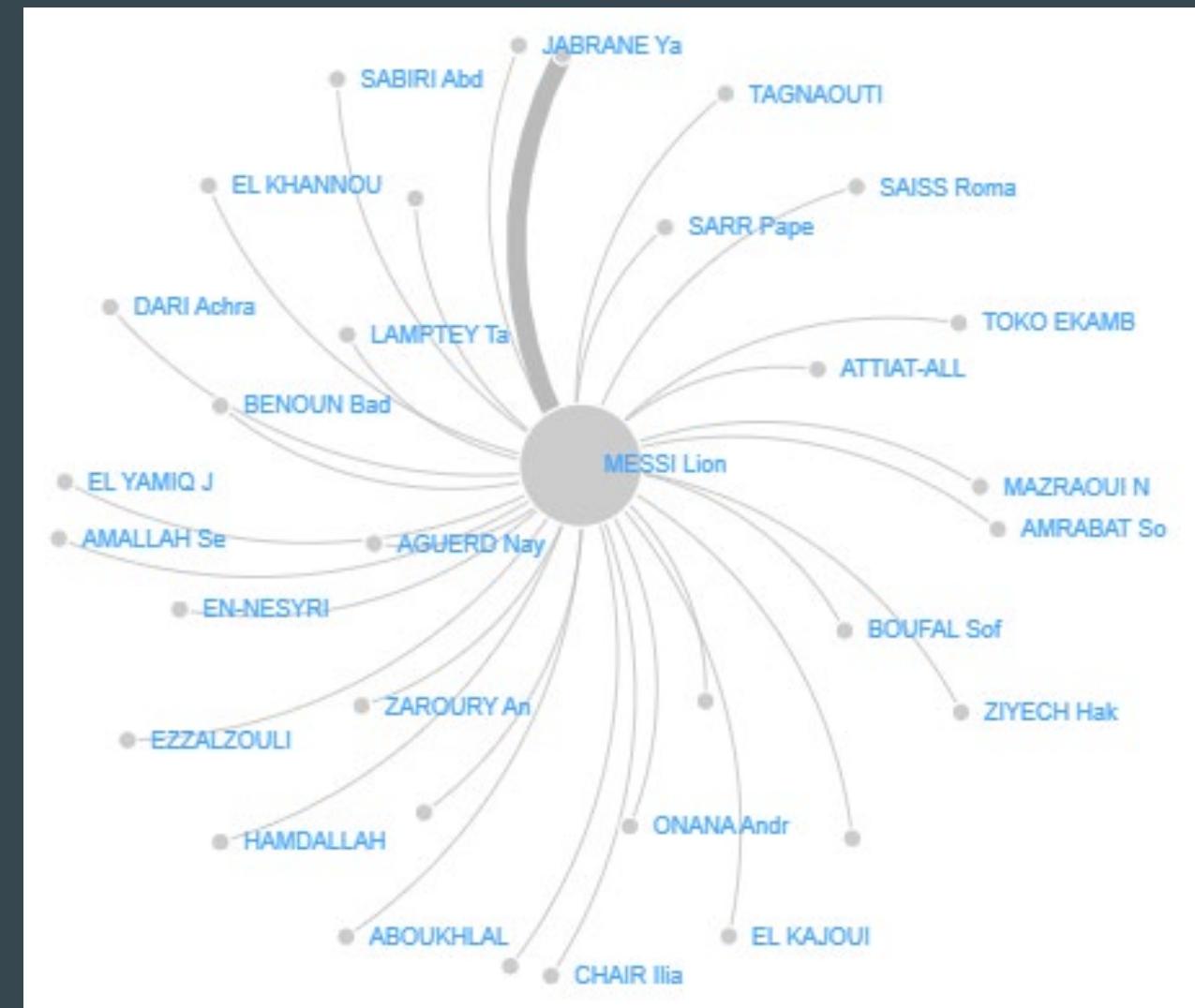
Combinando Dados Relacionais e Grafos

- Exemplo

```
model  
o  
grafo  
s  
  
WITH cteGrafo AS  
(  
    SELECT .....  
    FROM node.player NP1  
        , edge.playerXplayer FOR PATH AS EPP  
        , node.player FOR PATH AS NP2  
    WHERE MATCH(SHORTEST_PATH(NP1(-(EPP)->NP2)+))  
)  
SELECT g.* , c.Continent  
FROM cteGrafo g  
    INNER JOIN relational.Country c ON g.Country2 = c.countryID  
  
model relacional
```

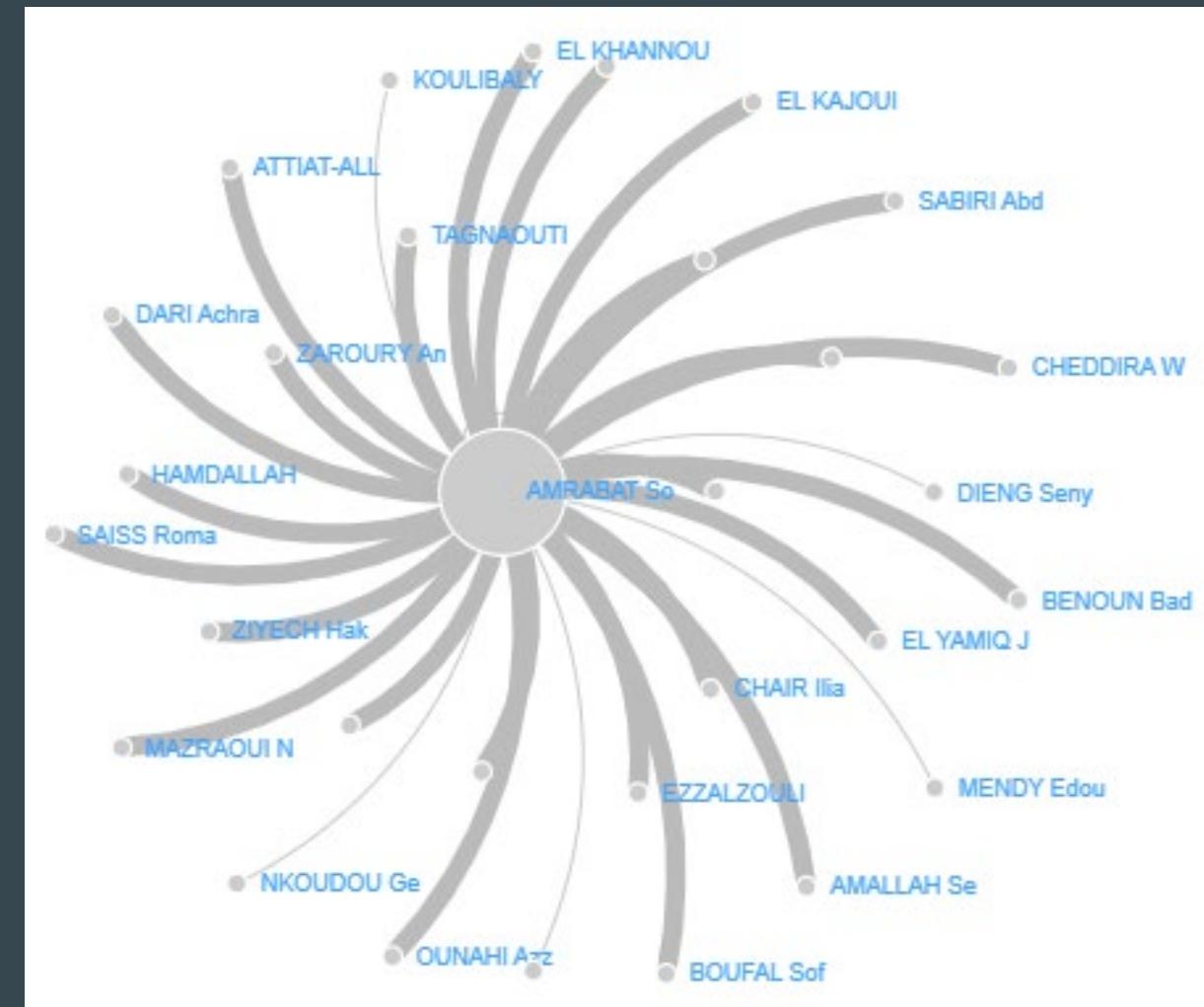
Visualização dos Grafos – POWER BI

- **Atletas africanos com relação de níveis 1 ou 2 com o atleta LIONEL MESSI**



Visualização dos Grafos – POWER BI

- Atletas africanos com relação de níveis 1 ou 2 com o atleta **AMRABAT (MARROCOS)**





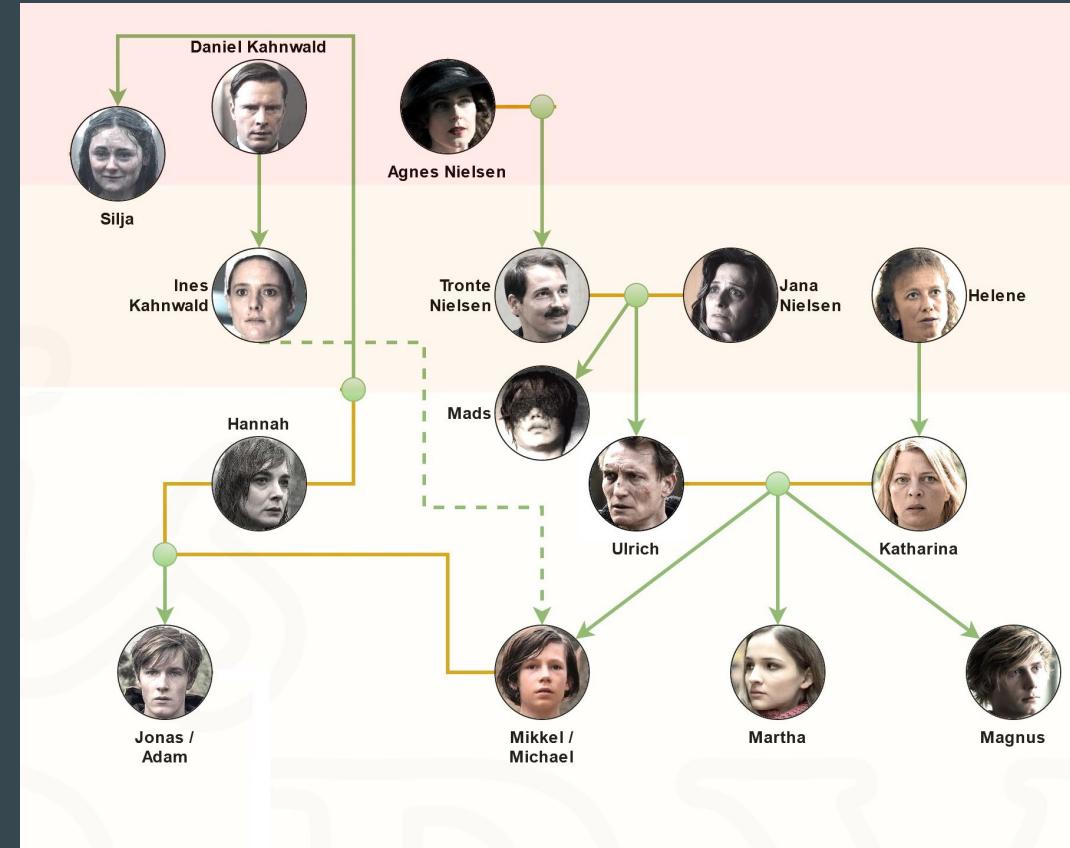
Vantagens dos Grafos



- Consultas complexas são muito mais fáceis de criar
- A complexidade da consulta não se altera com o número de níveis
- Performance nestes casos é muito melhor
- Pode-se **combinar dados** de tabelas de grafos com tabelas relacionais
- Integração com **POWER BI**

Comentários Finais

- Grafos são complementares ao modelo relacional
- Sintaxe TSQL é fácil de lidar
- Carga de dados não é elementar
- Escolha relacionamentos relevantes
- POWER BI oferece recursos para apresentação dos grafos





Leituras Sugeridas

1. Bate-bola com Grafos e SQL Server Parte 1 – Definindo o Modelo (bit.ly/sqlgraph1)
2. Bate-bola com Grafos e SQL Server Parte 2 – Implementando o Modelo (bit.ly/sqlgraph2)
3. Bate-bola com Grafos e SQL Server Parte 3 – Visualizando Dados (bit.ly/sqlgraph3)
4. Análise de Grafos no Azure SQL DB (bit.ly/sqlgraph4)
5. Material da palestra (<https://github.com/wcrivelini/articles>)



Perguntas

VIX25

www.dataes.com.br 

A black and white aerial photograph of a coastal city, likely Rio de Janeiro, Brazil. The image shows a large body of water in the foreground, a bridge spanning across it, and a dense urban area built on hills in the background. The sky is overcast.

Obrigado!



<https://www.linkedin.com/in/wagner-crivelini/>