Neo4j - Importação de Dados

Igor Avila Pereira igor.pereira@riogrande.ifrs.edu.br

Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS) Câmpus Rio Grande Divisão de Computação

Agenda

- Importando dados de um arquivo CSV
 - Arquivo CSV local
 - Arquivo CSV Remoto
 - Arquivo CSV com Cabeçalho
 - Arquivo CSV com Delimitador Personalizado
 - Importando Dados com Caracteres de Escape
 - Usando linenumber() com LOAD CSV
- Trabalhando com Grandes Quantidades de Dados
 - Definindo o número de linhas
- 3 Usando file() com LOAD CSV

Importando dados de um arquivo CSV

- O arquivo CSV a ser usado deve ter as seguintes características:
 - a codificação de caracteres deve ser UTF-8;
 - a terminação da linha final é dependente do sistema, por exemplo, é \n no Unix ou \r\n no Windows;
 - o terminador de campo padrão é vírgula;
 - Strings entre aspas são permitidas no arquivo CSV e as aspas são eliminadas ao ler os dados;
 - o caractere para citação de String é aspas duplas;
 - uma aspa dupla deve estar em uma String entre aspas e ter escape, seja com o caractere de escape ou uma segunda aspa dupla.

Query Cypher

Arquivo CSV local Arquivo CSV Remoto Arquivo CSV com Cabeçalho Arquivo CSV com Delimitador Personalizado Importando Dados com Caracteres de Escape Usando linenumber() com LOAD CSV

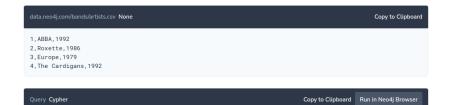
Arquivo CSV local

```
1,ABBA,1992
2,Roxette,1986
3,Europe,1979
4,The Cardigans,1992
```

```
LOAD CSV FROM 'file:///artists.csv' AS line
CREATE (:Artist {name: line[1], year: toInteger(line[2])})
```

Arquivo CSV Remoto

vear: toInteger(line[2])})



 Obs: Isso pode ser aplicado a todas os comandos de importação

LOAD CSV FROM 'https://data.neo4i.com/bands/artists.csv' AS line CREATE (:Artist {name: line[1].

Arquivo CSV com Cabeçalho

 Quando seu arquivo CSV tem cabeçalhos, você pode visualizar cada linha do arquivo como sendo um objeto.



Arquivo CSV com Delimitador Personalizado

- Às vezes, seu arquivo CSV tem outros delimitadores de campo além de vírgulas.
- Você pode especificar qual delimitador seu arquivo usa, usando FIELDTERMINATOR.
- A representação hexadecimal da codificação de caracteres unicode pode ser usada se precedida por \u.



Importando dados com caracteres de escape

 Neste exemplo, ambos temos aspas adicionais em torno dos valores, bem como aspas com escape dentro de um valor



Usando linenumber() com LOAD CSV

 Para determinados cenários, como depurar um problema com um arquivo csv, pode ser útil obter o número da linha atual em que LOAD CSV está operando. A função linenumber() fornece exatamente isso ou null



Trabalhando com Grandes Quantidades de Dados

 Se o arquivo CSV contiver um número significativo de linhas (aproximando-se de centenas de milhares ou milhões), CALL IN TRANSACTIONS pode ser usado para instruir o Neo4j a confirmar uma transação

```
Query Cypher

Copy to Clipboard

Run in Neo4j Browser

LOAD CSV FROM 'file:///artists.csv' AS line

CALL {

WITH line

CREATE (:Artist {name: line[1], year: toInteger(line[2])})

} IN TRANSACTIONS
```

Definindo o número de linhas

 Você pode definir o número de linhas como no exemplo, onde é definido como 500 linhas

```
Query Cypher

Copy to Clipboard

Run in Neo4j Browser

LOAD CSV FROM 'file:///artists.csv' AS line

CALL {

WITH line

CREATE (:Artist {name: line[1], year: toInteger(line[2])})
} IN TRANSACTIONS OF 500 ROWS
```

Usando file() com LOAD CSV

artists.csv None Copy to Clipboard 1, ABBA, 1992 2, Roxette, 1986 3, Europe, 1979 4, The Cardigans, 1992

Query Cypher Copy to Clipboard Run in Neo4j Browser LOAD CSV FROM 'file:///artists.csv' AS line RETURN DISTINCT file() AS path

Neo4j - Importação de Dados

Igor Avila Pereira igor.pereira@riogrande.ifrs.edu.br

Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS) Câmpus Rio Grande Divisão de Computação