

Практическа работа 2

Самолетни трасета

Ние желаем да конструираме с базата данни Neo4j интернет система за оптимизиране на пътническите полети. За целта ще използваме наличните свободни данни от този сайт:

<https://openflights.org/data>

Необходимите файлове са само два:

airports.dat

routes.dat

Тези файлове трябва да се изчистят и напаснат с език по ваш избор, така че да има съответствие между съществуващите летища и съществуващите полети.

Ако изпитвате трудности с подготовката на данните, то при лично запитване могат да бъдат предоставени вече почистени и напаснати такива на цената на -30 точки от крайната оценка.

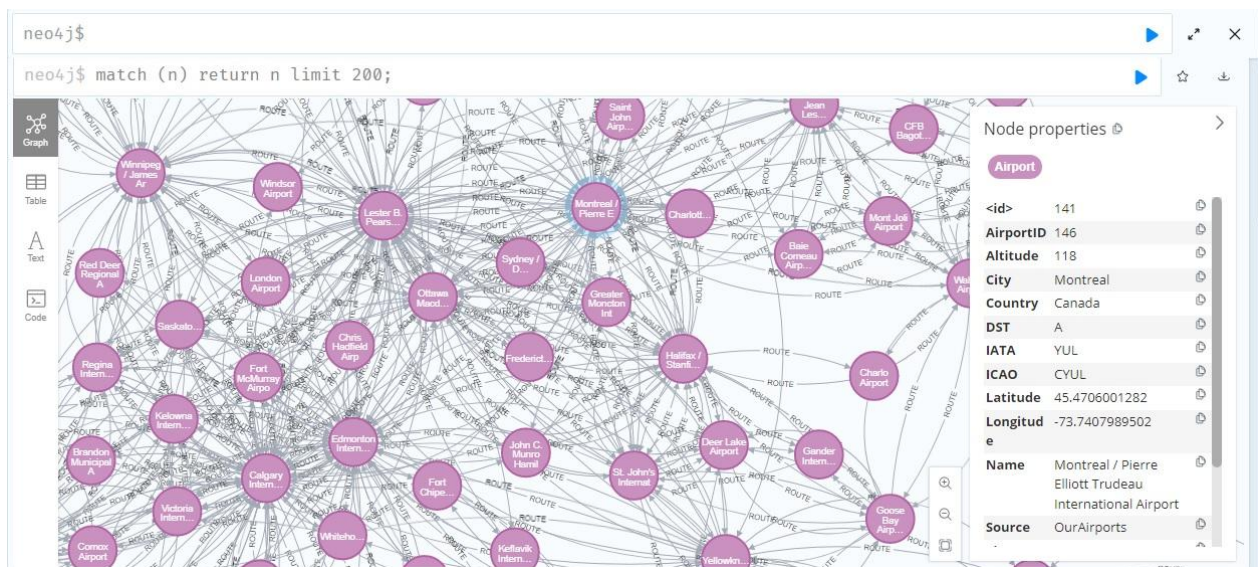
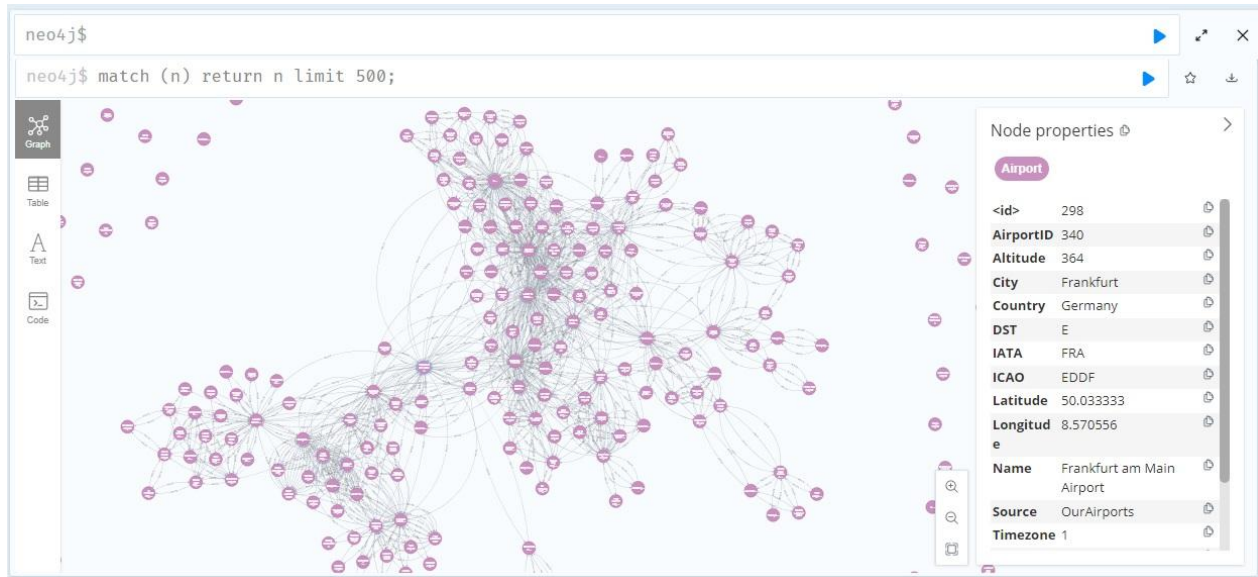
Схемата на създадената база данни трябва да отговаря на следната схема:

AIRPORT-[ROUTE]-AIRPORT

Всяко летище трябва да съдържа всички параметри от файла с данни. Всеки път трябва да съдържа всички параметри от файла с пътищата.

Зареждането на двата файла в базата данни трябва да стане с командата `call apoc.import.csv()`, като при всяко друго решение тук ще бъдат отнемани -10 точки.

Файловете за летищата и пътищата са вече в необходимият формат, но трябва да им се добави мапинг за да се получи необходимото зареждане по начина, който е показан по-долу:



file	source	format	nodes	relationships	properties	time	rows	batch
1	"progress.csv"	"file"	"csv"	6073	66934	553356	30523	0

Броят на нодовете и връзките могат да варират в зависимост от текущите условия, но не с много!

Въпроси за решаване:

- 1.Кои летища можем да посетим от София и които са на 1 дъга разстояние?
- 2.А на 2 дъги разстояние?
- 3.Същото за Пловдив
- 4.Най-краткият път от Пловдив до Сидни?
- 5.А най-бързият?
- 6.Същото, но за Лондон
- 7.Същото от Монтреал до София
- 8.Най-натовареното летище?

Освен решенията е необходимо да се направи и кратко представяне на вашият подход и резултати под формата на презентация, която ще бъде показана и обяснена пред колегите в последният 7-ми курс, предназначен точно за това.

Оценяване:

Крайната максимална оценка е 100%.

Доставени почистени файлове персонално -30%.

Използван друг начин за зареждане на нодове и зависимости -10%.

Заредени само половината или по-малко нодове и връзки -10%.

Липса на връзки -10%

Грешна заявка -10%

Крайният срок е последният 7-ми курс, на който резултатите трябва да се представят пред участниците и депозират в мудъл.

Непредставени работи получават 0%.

Приятна работа!