SAPIENZA
UNIVERSITÀ DI ROMA

Lorenzo Lancia
Alessandro Gallo

Master Degree in Data Science
Sapienza Università di Roma

AtacNoSql

Graph Database
Project 3 for DMDS course

Indice

Introduction

- Aim of this project was to build a No-Sql database using the same dataset of the previous projects
- So again, we used open data about public transport network of city of Rome provided by ATAC.
  `http://www.agenziamobilita.roma.it/it/progetti/open-data/dataset.html`
- The DBMS used is Neo4J[1]

[1] Ver 3.0.0

Open Data

ATAC provides the data in the GTFS format.

What is GTFS?

The General Transit Feed Specification (GTFS) defines a common format for public transportation schedules and associated geographic information. [2]

Unfourtunatly not all tables required in GTFS standard are available in the ATAC dataset.

[2] `https://developers.google.com/transit/gtfs/#how-do-i-start`

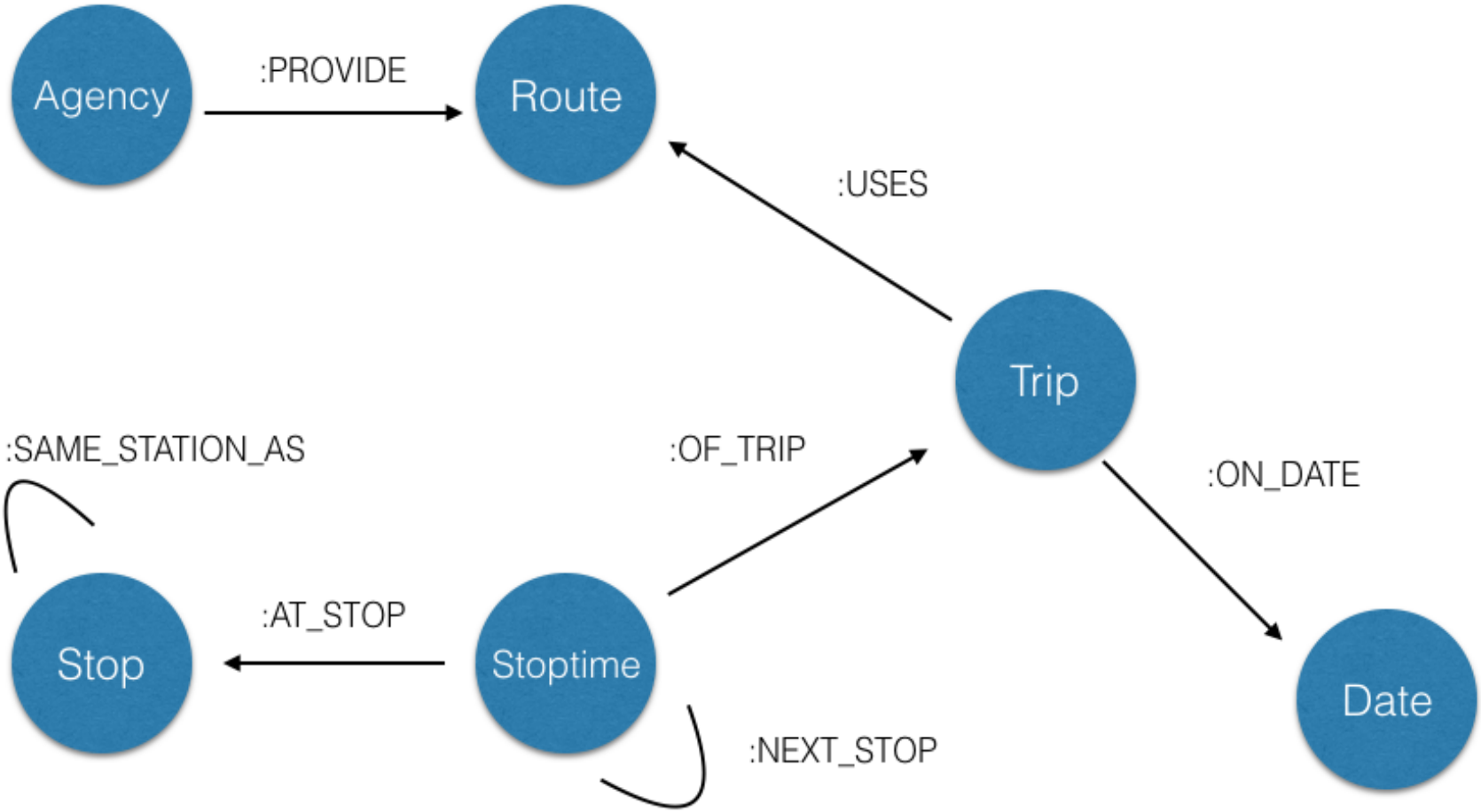
File Used

`stops.txt`	Individual locations where vehicles pick up or drop off passengers.
`routes.txt`	Transit routes. A route is a group of trips that are displayed to riders as a single service.
`trips.txt`	Trips for each route. A trip is a sequence of two or more stops that occurs at specific time.
`stop_times.txt`	Times that a vehicle arrives at and departs from individual stops for each trip.
`calendar_dates.txt`	Exceptions for the service IDs defined in the `calendar.txt` file. If `calendar_dates.txt` includes ALL dates of service, this file may be specified instead of `calendar.txt`.

Table: Data Set files

SAPIENZA
UNIVERSITÀ DI ROMA

The Graph



Query

1) List all routes that stops at "Policlinico"

```
match
(s:Stop)<-[:AT_STOP]-()-[:OF_TRIP]->()-[:USES]->(r:Route
    )
where s.name="POLICLINICO"
or s.name ="POLICLINICO "
return distinct r.id;
```

results: N2L N13 490 495 61 N2 649 2 19 3 N10 88 N11 MEB MEB1

2) List all routes that a stop at a stop containing the word "DE LOLLIS"

```
match
(s:Stop )<-[:AT_STOP]-()-[:
    OF_TRIP]->()-[:USES]->(
    res:Route)
where
s.name CONTAINS "DE LOLLIS"
return distinct res.id;
```

results:

```
   492
N10
C3
C2
     2
    19
     3
    71
```

3) List the stops of the autobus 445

```
match (r:Route)<-[:USES]-()
    <-[:OF_TRIP]-()-[:
    AT_STOP]->(fermata)
where r.id = "445"
return DISTINCT fermata.name
    ;
```

```
BOLOGNA
VENTUNO APRILE- VILLA RICOTTI
VENTUNO APRILE- NARDINI
LANCIANI- BOLDETTI
LANCIANI- DE PETRA
MONTI TIBURTINI- NOMENTANA
MONTI DI PIETRALATA
CURIONI- DE LORENZO
CURIONI- COLLINA LANCIANI
CURIONI
CURIONI- PENTA
CURIONI- REPOSSI
LARGO LANCIANI
LANCIANI- WINCKELMANN
VENTUNO APRILE- RICOTTI
CARACI- MINISTERO INFRASTRUTTURE T
```

SAPIENZA UNIVERSITÀ DI ROMA

4) How many trains of MEB runs in a day?

```
match
(date :Date)<-[:ON_DATE]-(aa
    )-[:USES]->(r:Route {id:
     "MEB"})
return count(aa), date.date
order by date.date limit 7;
```

```
count(aa) | date.date
----------+----------
      375 |  20160314
      375 |  20160315
      375 |  20160316
      375 |  20160317
      400 |  20160318
      267 |  20160319
      241 |  20160320
```

SAPIENZA UNIVERSITÀ DI ROMA

5) Which line departs from “VERANO” after 5pm

```
match
(s:Stop {name: "VERANO"})
    <-[:AT_STOP]-(time:
    Stoptime)
-[:OF_TRIP]->()-[:USES]->(r)
where time.departure_time> "
    17:00:00"
return r.id
order by time.departure_time
     asc
 limit 5;
```

```
r.id
----
 545
 542
 163
   3
  71
```

SAPIENZA UNIVERSITÀ DI ROMA

6) Which are the three most frequented stops on sunday?

```
match (d:Date {date:"
    20160320"})<-[:ON_DATE
    ]-(trips:Trip)
<-[:OF_TRIP]-()-[:AT_STOP
    ]->(s:Stop)
 return count(trips.id) as
     cnt, s.name
 order by cnt desc
 limit 3 ;
```

```
cnt   | s.name
------+---------------
3395  | TERMINI
2770  | PIAZZA VENEZIA
2493  | CONCA D'ORO
```

SAPIENZA UNIVERSITÀ DI ROMA

7) List all of stops served by nocturne bus

```
match
(r:Route)<-[:USES]-()<-[:
    OF_TRIP]-()-[:AT_STOP
    ]->(s:Stop)
 where
 r.id STARTS WITH "N"
 return DISTINCT s.name, r.
     id  limit 4;
```

```
s.name                    | r.id
--------------------------+-----
PIAZZA SEMPIONE           | N22
MENENIO AGRIPPA           | N22
NOMENTANA- RIONERO        | N22
NOMENTANA- SANNAZZARO     | N22
```

SAPIENZA UNIVERSITÀ DI ROMA

8) List all routes that connects Rebibbia to Tiburtina Station

```
match (s:Stop )<-[:AT_STOP
    ]-()-[:NEXT_STOP*]->
(l)-[:AT_STOP]->(fs:Stop),
(l)-[:OF_TRIP]->()-[:USES
    ]->(r:Route)
where s.name CONTAINS "
    REBIBBIA" and (
  fs.name="TIBURTINA" or
  fs.name CONTAINS  "
      STAZIONE TIBURTINA")
return distinct r.id;
```

```
r.id
----
 163
N23
N2
120F
MEB
```