

TD 3 : BD avancées

Année 2013-2014

Modélisation : Etude de cas

La RATP participe à une action “Open Data” et met en ligne des données sur son réseau. Le but de cet exercice est de modéliser les données fournies par la RATP à l’aide d’un schéma entité association, dans le but de comprendre ce que représentent les données, et de compléter le schéma fourni avec les clefs nécessaires.

Les données sont dans le format GTFS. “Le General Transit Feed Specification est un format de données développé par Google afin de promouvoir la compatibilité et l’interopérabilité des services dédiés aux transports.”

- Une période de service est soit une date, soit un intervalle de dates.
- Une route correspond à un trajet de bus, de métro, de tram, de RER.
- Une course (“trip”) est une séquence d’arrêts avec un horaire pour chaque arrêt.

Pour préciser les horaires, on doit associer aux courses une période et un trajet (“route”). On connaît également les temps de transfert à pied entre les arrêts. En pratique, les données sont réparties dans huit fichiers au format CSV. Nous n’utiliserons que sept de ces fichiers, et un sous-ensemble des attributs fournis.

```
==> calendar.txt <==
```

debut et fin de chaque periode de service

```
service_id,start_date,end_date
```

```
11396973,20130707,20130901
```

```
==> calendar_dates.txt <==
```

dates precises de chaque service

```
service_id,date
```

```
1364400,20130401
```

```
==> routes.txt <==
```

numero, nom et direction de chaque route (trajet).

```
route_id,route_short_name,route_long_name
```

```
918044,"1","(CHATEAU DE VINCENNES <-> LA DEFENSE) - Aller"
```

```
918045,"1","(CHATEAU DE VINCENNES <-> LA DEFENSE) - Retour"
```

```
==> stop_times.txt <==
```

horaires d’arrivee et de depart pour chaque course ("trip")

```
trip_id,arrival_time,departure_time,stop_id,stop_sequence
```

```
1013644000942075,16:10:00,16:10:00,4025388,1
```

```
1013644000942075,16:11:00,16:11:00,4025390,2
```

```
1013644000942075,16:12:00,16:12:00,4025392,3
```

```
==> stops.txt <==
```

code, nom, adresse, longitude et latitude des arrêts

```
stop_id,stop_name,stop_desc,stop_lat,stop_lon
```

```
4025457,,"PALAIS ROYAL - COMEDIE FRANCAISE","155-157 RUE SAINT HONORE - 75101",48.8629200485678,2
```

```
4025456,,"MUSEE DU LOUVRE","PLACE DU CARROUSEL - 75101",48.86119444420634,2.3345576627019695
```

```
==> transfers.txt <==
```

correspondances et temps de correspondance (en secondes)

```
from_stop_id,to_stop_id,min_transfer_time
```

```
4024670,4025440,19
```

```
==> trips.txt <==
associe une course ("trip") a une route et un service
Pour un route_id et un service_id, on associe plusieurs trip_id (Il y a plusieurs
courses / jour / route)
route_id,service_id,trip_id,trip_short_name,direction_id
990303,1388894,13913888940877947,139,0
```

En vous regroupant en petits groupes (3-5 personnes), répondez aux questions suivantes.

1. Quelles sont les entités ? Y a-t-il des entités faibles ? De l'héritage ?
2. Quelles sont les associations ?
3. Donnez les cardinalités de chaque association.
4. Donnez les clefs candidates pour entité.