



M1 INFORMATIQUE AIGLE

**HMIN122M**

MINI-PROJET : ENTREPÔTS DE DONNÉES

---

## Rendu sur la modélisation d'un entrepôts de données

---

Bachar RIMA

Joseph SABA

Tasnim SHAQURA MUHAMMAD

Jérémy BOURGIN

23 octobre 2018

# Table des matières

|   |          |
|---|----------|
| <b>Introduction</b>                                   | <b>2</b> |
| <b>Questions</b>                                      | <b>3</b> |
| Question 1 . . . . .                                  | 3        |
| Question 2 . . . . .                                  | 3        |
| Question 3 . . . . .                                  | 4        |
| Question 4 . . . . .                                  | 4        |
| Questions 5 et 6 . . . . .                            | 4        |
| <i>Data Mart</i> de « voyages » . . . . .             | 4        |
| <i>Data Mart</i> de « maintenance » . . . . .         | 5        |
| <i>Data warehouse</i> de <i>tam-voyages</i> . . . . . | 6        |
| Question 7 . . . . .                                  | 7        |
| Question 8 . . . . .                                  | 7        |

# Introduction

Dans le cadre du mini-projet du module **HMIN122M**, nous avons décidé de modéliser un entrepôts de données pour le réseau de transport public de Montpellier, *tam-voyages*. Pour ce faire, nous avons proposé des *data marts* formant le *data warehouse* et permettant de réaliser des requêtes analytiques sur un ensemble important de données. Cette modélisation permettra ainsi de mettre en œuvre un outil d'analyse permettant de bien répondre aux problématiques suivantes :

1. Comment *tam-voyages* pourront-ils augmenter leur taux de ventes en se basant sur la circulation du réseau<sup>1</sup> ?
2. Comment *tam-voyages* pourront-ils suivre l'évolution et la maintenance de leurs matériaux de manière à réduire les dépenses qui y sont associées ?

Ces problématiques seront ainsi adressées en analysant les actions et opérations effectuées par *tam-voyages*, notamment en choisissant celles qui paraissent les plus pertinentes et les plus importantes en termes de données intégrées et flexibilité des critères d'analyse.

---

1. en particulier en examinant les lignes de tramway et les bus

# Questions

## Question 1

Les actions/opérations effectuées par *tam-voyages* considérées :

- Les voyages.
- La maintenance de véhicules.
- Les ventes de tickets et les abonnements.
- Les amendes.

## Question 2

1. exemples de requêtes analytiques pour l'action « voyages » :
  - le nombre de voyageurs par bus, utilisant des tickets pour le mois de juillet.
  - le prix moyen par type de ticket pour chaque voyage pendant les vacances de Noël de 2018.
  - le nombre de voyageurs abonnés par ligne pour chaque voyage pour les deux derniers mois.
  - l'arrêt le plus fréquenté par toutes les lignes de circulation.
2. exemples de requêtes analytiques pour l'action « maintenance » :
  - le nombre de bus maintenus pour le mois de septembre 2018.
  - les X premières lignes ayant le nombre maximale de maintenances par mois.
  - les X premières véhicules nécessitant le plus de maintenance pour les 6 dernier mois.
3. exemples de requêtes analytiques pour l'action « ventes » :
  - le nombre d'abonnés ayant plus que 26 ans pour le mois d'août 2018.
  - le nombre d'abonnés par date de naissance pour l'année 2018.
  - les types d'abonnement les plus fréquents pour l'année 2018.
4. exemples de requêtes analytiques pour l'action « amendes » :

- les lignes qui ont générées le plus d’amendes pour les deux derniers mois.
- les lignes les plus contrôllées de la semaine dernière.
- le nombre des abonnés qui ont reçu des amendes par ligne, l’avant-midi.
- la somme total d’amendes rapportée par type de voyageur par ligne pour le dernier mois.

### Question 3

Les actions considérées, par ordre d’importance :

1. « voyages ».
2. « ventes ».
3. « maintenance ».
4. « amendes ».

### Question 4

Les actions les plus pertinentes à analyser vis-à-vis les problématiques avancées sont « voyages » et « maintenance » qu’on traitera de la manière suivante :

**voyages** : modèle en étoile détaillé.

**maintenance** : modèle en étoile *moins* détaillé, en particulier le modèle intitulé "*periodic snapshot*".

### Questions 5 et 6

#### *Data Mart* de « voyages »

Les mesures de la table des voyages sont :

- `travel_price` : additive

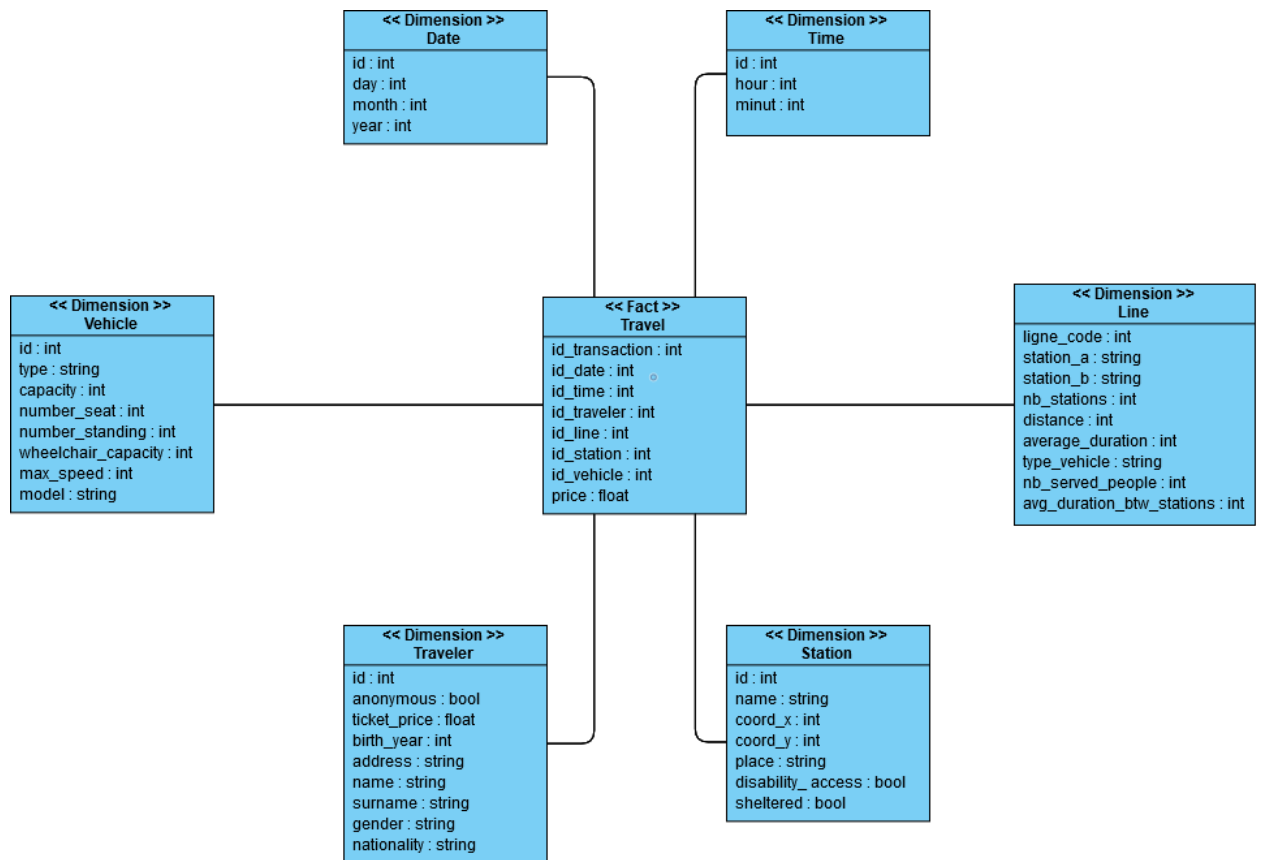


FIGURE 1 – modèle en étoile de l'action « voyages »

### *Data Mart* de « maintenance »

Les mesures de la table des maintenances sont :

- `cost` : additive
- `estimated_time` : additive

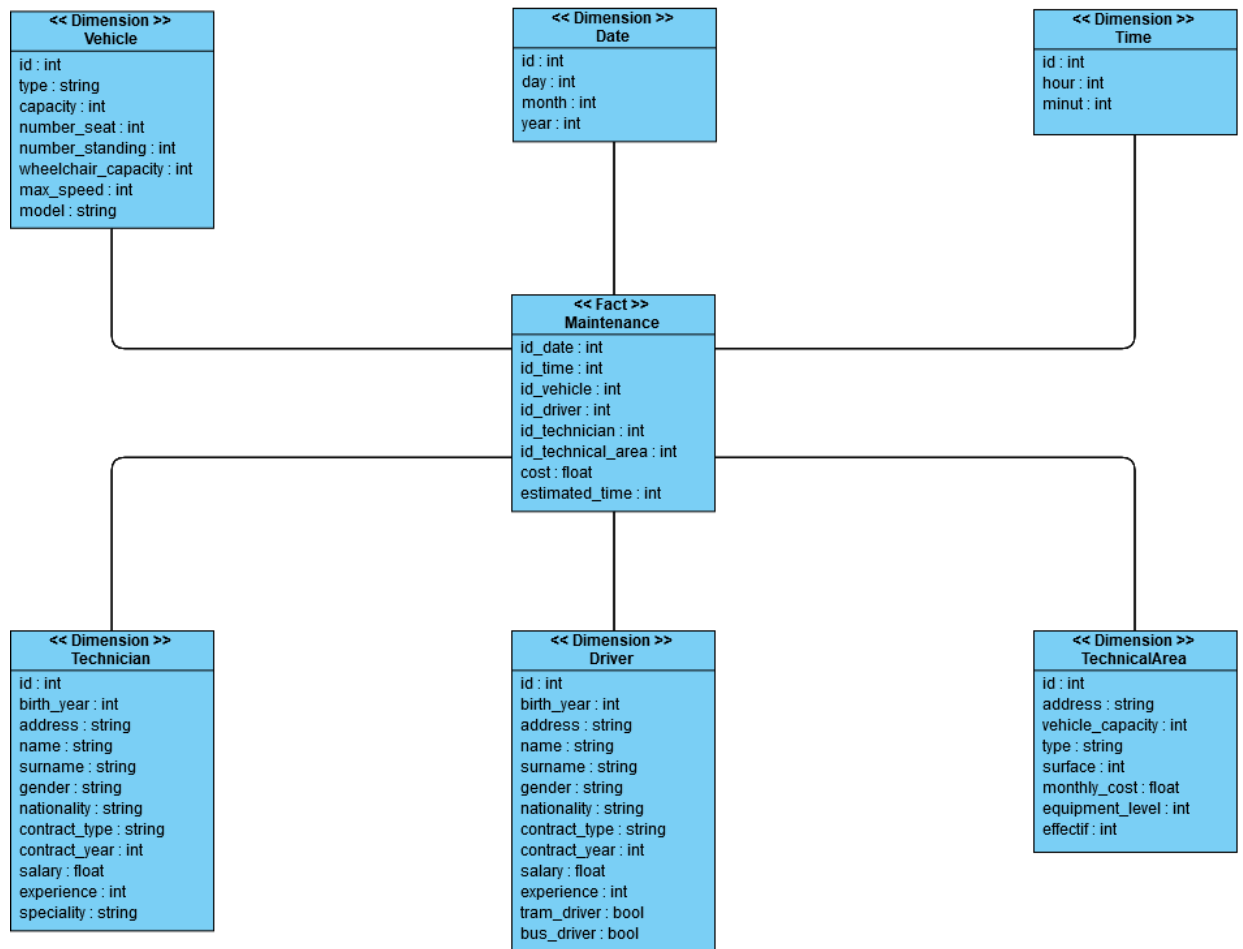


FIGURE 2 – modèle en étoile de l'action « maintenance »

*Data warehouse de tam-voyages*

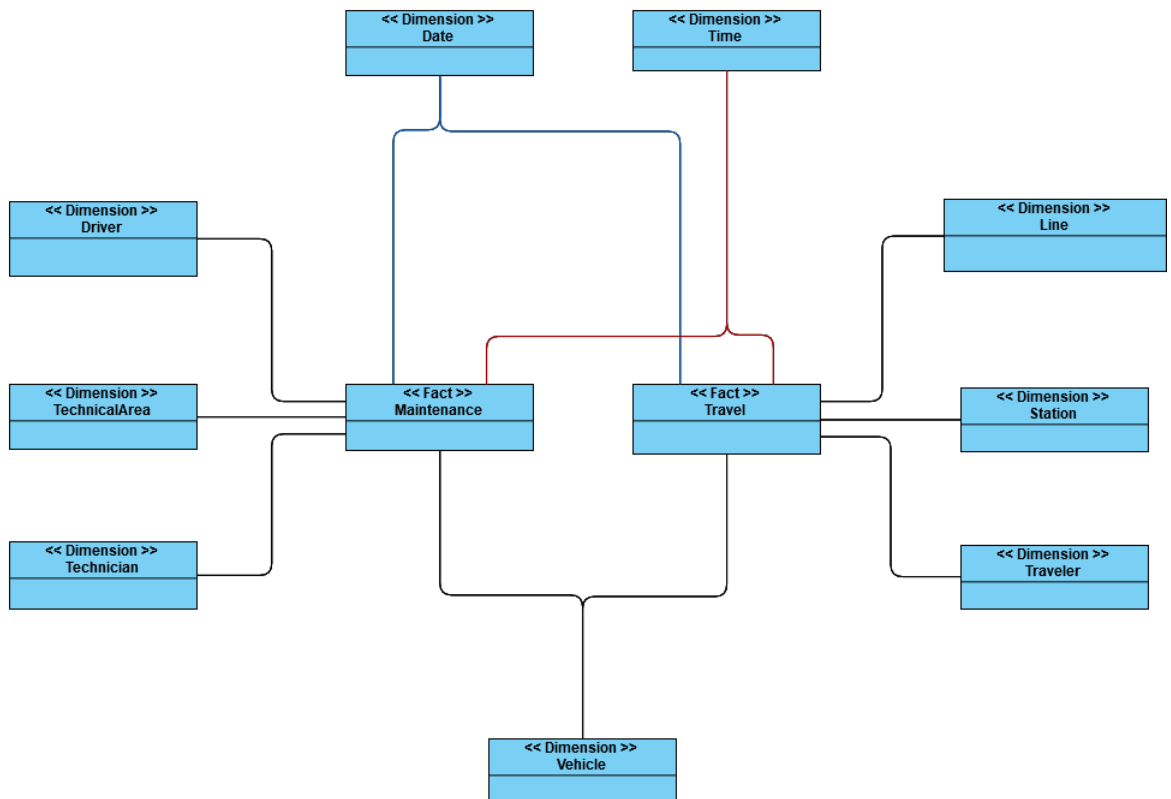


FIGURE 3 – le *data warehouse* résultant

Question 7

Question 8

| <u>id_travel</u> | <u>id_date</u> | <u>id_time</u> | <u>id_traveler</u> | <u>id_line</u> | <u>id_station</u> | <u>id_vehicle</u> | <u>price_travel</u> |
|------------------|----------------|----------------|--------------------|----------------|-------------------|-------------------|---------------------|
|------------------|----------------|----------------|--------------------|----------------|-------------------|-------------------|---------------------|

TABLE 1 – Travel

| id_line | nb_line | start_station        | end_station  | nb_stations | distance_travel | avg_duration_travel | type_vehicle | nb_served_people | avg_duration_btw_stations |
|---------|---------|----------------------|--------------|-------------|-----------------|---------------------|--------------|------------------|---------------------------|
| 101     | 7       | Hôtel du département | La Martelle  | 34          | 1510            | 45                  | bus          | 40               | 5                         |
| 102     | 7       | Hôtel du département | Les Bouisses | 34          | 1515            | 45                  | bus          | 40               | 4                         |
| 201     | 1       | Odysseum             | Mosson       | 30          | 15700           | 50                  | tram         | 150              | 2                         |

TABLE 2 – Line



| id_station | name                 | coord_x   | coord_y  | place              | disability_access | sheltered_for_rain |
|------------|----------------------|-----------|----------|--------------------|-------------------|--------------------|
| 1          | Hôtel du département | 43.622108 | 3.835276 | Avenue des moulins | <i>false</i>      | <i>true</i>        |
| 2          | Pergola              | 43.617558 | 3.839687 | Rue Paul Rimbaud   | <i>false</i>      | <i>true</i>        |

TABLE 3 – Station

| id_vehicle | type | model              | capacity | max_speed | nb_seats | standing_capacity | wheelchair_capacity |
|------------|------|--------------------|----------|-----------|----------|-------------------|---------------------|
| 1          | bus  | TransBus Enviro30  | 40       | 200       | 20       | 40                | 5                   |
| 2          | tram | Alstom Citadis 401 | 300      | 300       | 150      | 150               | 20                  |

TABLE 4 – Vehicle