

**Universités de Montpellier**  
**Faculté des Sciences**  
**Département d'informatique**

HAI708I

Entrepôt de données et big data

**PROJET**

**HERTZ**



**Réalisé Par**

**PROKOPY Lucie**

**OTTE Arthur**

**MEFLAH Wided**

**RAMDANE Souhaib**

**M1 GL**

## **Introduction**

Le monde connu un véritable changement dans les habitudes de consommation des citoyens, notamment dans le secteur de l'automobile où devenir propriétaire d'un véhicule n'est plus nécessairement une nécessité. La location de voiture est un marché qui génère une demande de plus en plus forte dans le monde entier. En effet, c'est pour les particuliers comme pour les professionnels, les voyageurs souhaitent des services simples, pratiques, mais surtout efficaces. De ce fait, les entreprises de location de voiture visent à l'analyse de leurs données afin de répondre à ces besoins. Dans le document présent, nous focalisons sur l'étude de l'entreprise HERTZ de location de voiture, en proposant une analyse, conception et implémentation d'un entrepôt de données permettant une meilleure prise de décision au sein de l'entreprise.

## **I. L'analyse**

### **1.1. Présentation et Objectifs de l'entreprise**

Hertz est une société américaine créée en 1918 par John Daniel Hertz. Elle est considérée comme un leader du marché de la location automobile avec une présence dans plus de 150 pays, 10 000 agences dont 220 en Europe.

Le but principal de l'entreprise est de maximiser les revenus et de gagner le plus d'argent possible. Pour le réaliser il faut garder les clients courants, et en gagner de nouveaux, mais aussi cultiver la confiance et la fidélité des clients.

Maintenant, Hertz doit se récupérer de la pandémie. Pour cela, l'entreprise doit assurer une bonne expérience et créer des avantages pour les clients fidèles, comme la collection des points et des statuts différents. Le but de cela est que les clients vont continuer de prêter les voitures chez Hertz, même s'il n'est pas vraiment nécessaire. Pour une bonne expérience, on doit aussi s'assurer que les voitures sont toujours en bon état. En particulier parce qu'il y a une concurrence forte.

Pour un profit maximal Hertz doit pronostiquer précisément la quantité de voitures nécessaire à un endroit particulier, ainsi que les modèles préférés.

Pour la gestion des véhicules, Hertz doit choisir une méthode: risquer d'avoir trop de voitures ou pas assez? S'il n'y a pas assez, ils risquent de perdre des clients à une autre entreprise,

mais avec trop de voitures ils perdent de l'argent. Garder les clients est probablement plus important, en ce moment.

Un autre but doit être plus écologiste et réduire ses émissions, mais au moment Hertz a seulement une collection verte qui était introduite en 2014.

## **1.2. Position de Hertz sur le marché**

L'entreprise Hertz est une des plus grandes et plus vieilles sur le marché de location de voiture. Actuellement, elle finit septième en ce qui concerne la taille d'entreprise. Elle emploie plus de 20.000 salariés à environ 10.000 endroits dans 160 pays et son revenu total annuel se chiffre par milliards. Néanmoins, Hertz avait de difficiles temps ces dernières années, comme la pandémie Covid-19 a fortement réduit le nombre de voyageurs et du coup les besoins de voitures de louage. De 2020 jusqu'à 2021, l'entreprise était en faillite. Depuis, elle s'est reposée un peu, mais il semble qu'il ne suffit pas encore: Jusqu'au début de 2022, ils avaient une promotion forte où chaque nouveau membre de leur programme de fidélité (Hertz Gold Plus Rewards) était automatiquement donné le statut élite (Five Star) pour une durée d'environ 6 mois. Le Five Star statut vient avec plus de points de bonus pour chaque location et plus de liberté en ce qui concerne le choix de voiture. Normalement, il est nécessaire de louer au moins 12 fois chez eux pendant une année pour qu'on le reçoive. La promotion montre qu'il est important pour Hertz de rapidement gagner de nouveaux clients. Elle encourage aussi de rester dans le Five Star statut après qu'elle n'a plus d'effet, donc de louer des voitures au moins 12 fois chaque année, ce qui aide à garder les clients très lucratifs.

## **1.3. Services et produits propose**

L'entreprise propose de nombreux services associés avec les véhicules automobiles. Premièrement, il y a les services de location qui comportent la location de voiture et voiture électrique, la location de véhicules utilitaires, la location de moto et la location d'auto-caravanes. Il y a encore la distinction entre location normale (pour plusieurs journées) et location à l'heure. En outre, on peut user de plusieurs accessoires comme des assurances supplémentaires, des équipements neige ou l'achat de carburant. Deuxièmement, Hertz vend une partie de ses voitures, mais seulement aux professionnels de la vente automobile.

## **1.4. Les revenu de Hertz**

Les formes de revenu de Hertz sont intrinsèquement liées à ses services. La source principale est la location de voiture dont les tarifs varient et sont influencés par plusieurs facteurs comme l'âge et statut du locataire, la date et durée de louage, le modèle de voiture et accessoires supplémentaires. En dehors de cela, une deuxième partie des revenus de Hertz vient de la vente de voitures aux professionnels.

### **1.5. Informations pour la prise de décision au sein de l'entreprise**

Pour que HERTZ optimise ses possibilités d'analyse de données et facilite sa prise de décisions, les informations suivantes sont utiles :

- Connaître les pays, villes et les agences avec le plus grand nombre de demandes de location.
- Connaître les modèles et les types (manuelle, automatique) de voitures les plus demandés pour la location.
- Connaître les périodes de l'année quand les locations sont populaires (périodes de tourisme).
- Connaître les services supplémentaires les plus ajoutés par les clients lors d'une réservation de voiture (assurance accident , protection contre le vol ...)
- Connaître les types de clients: âge? aisé? voyageant souvent? louent longtemps? fiable (retournent les voitures à temps et dans de bonnes conditions)? enfants?
- Connaître les bons endroits pour les agences dans une ville (bien accessibles sans voiture)
- Connaître les conditions des voitures (nombre de km roulés, pneus d'hiver, défaut ...)
- Connaître la régularité des réservations annulées
- Connaître les parcs où les voitures sont plus retournées en fin de journée. Combien de clients retournent la voiture dans une autre agence?
- Connaître l'effet des promotions: combien de clients est-ce qu'on a gagné? Devrait-on prolonger la promotion?
- Les voitures et accessoires suffisent-ils pour satisfaire les demandes?

### **2. Indiquez les actions / opérations à tracer pour récupérer ces informations.**

- La réservation / le paiement / la prise en main de voiture
- La distribution des voitures
- L'achat des voitures
- La maintenance des voitures

**3. Pour chaque action / opération, proposez au moins trois traitements possibles (i.e., requêtes analytiques) permettant d'aider à la prise de décision sur le sujet.**

La réservation / le paiement / la prise en main de voiture

Le nombre de réservations groupées par lieu pour une période définie

Le nombre de paiements effectués par type de client (degré de fidélité) pour une période donnée

Le montant total des paiements par modèle de voiture pour une période définie.

Distribution

Nombre total de voitures stockées dans un pays particuliers

Nombre total de voitures stockées par agence

Les agences qui stockent le plus grand nombre de voitures.

Achat de voiture

Montant d'achat de voiture d'une année

Montant d'achat de voiture d'une année groupé par modèle

Quantités des voitures électriques achetées par an

La maintenance des voitures

La somme des dépenses de maintenance par modèle de véhicule par année

La somme des dépenses de maintenance par agence pour une période donnée

Montant total des maintenances effectuées pour une année.

**4. Ordonnez les actions par ordre d'importance / rentabilité potentielle (e.g. augmentation des ventes vs. utilisation optimale de l'espace de stockage dans le magasin).**

1. La réservation / le paiement / la prise en main de voiture
2. La maintenance des voitures
3. La distribution
4. L'achat des voitures

**5. Identifiez les deux actions / opérations les plus importantes à analyser.**

1. la réservation / le paiement / la prise en main de voiture

## 2. la maintenance des voitures

**6. Pour chaque actions / opérations, concevez un data-mart indépendant (c'est à dire, un modèle en étoile)**

**a. Pour l'action / opération la plus importante, prévoyez une modélisation détaillée dans la table des faits.**

### **La réservation / le paiement / la prise en main de voiture:**

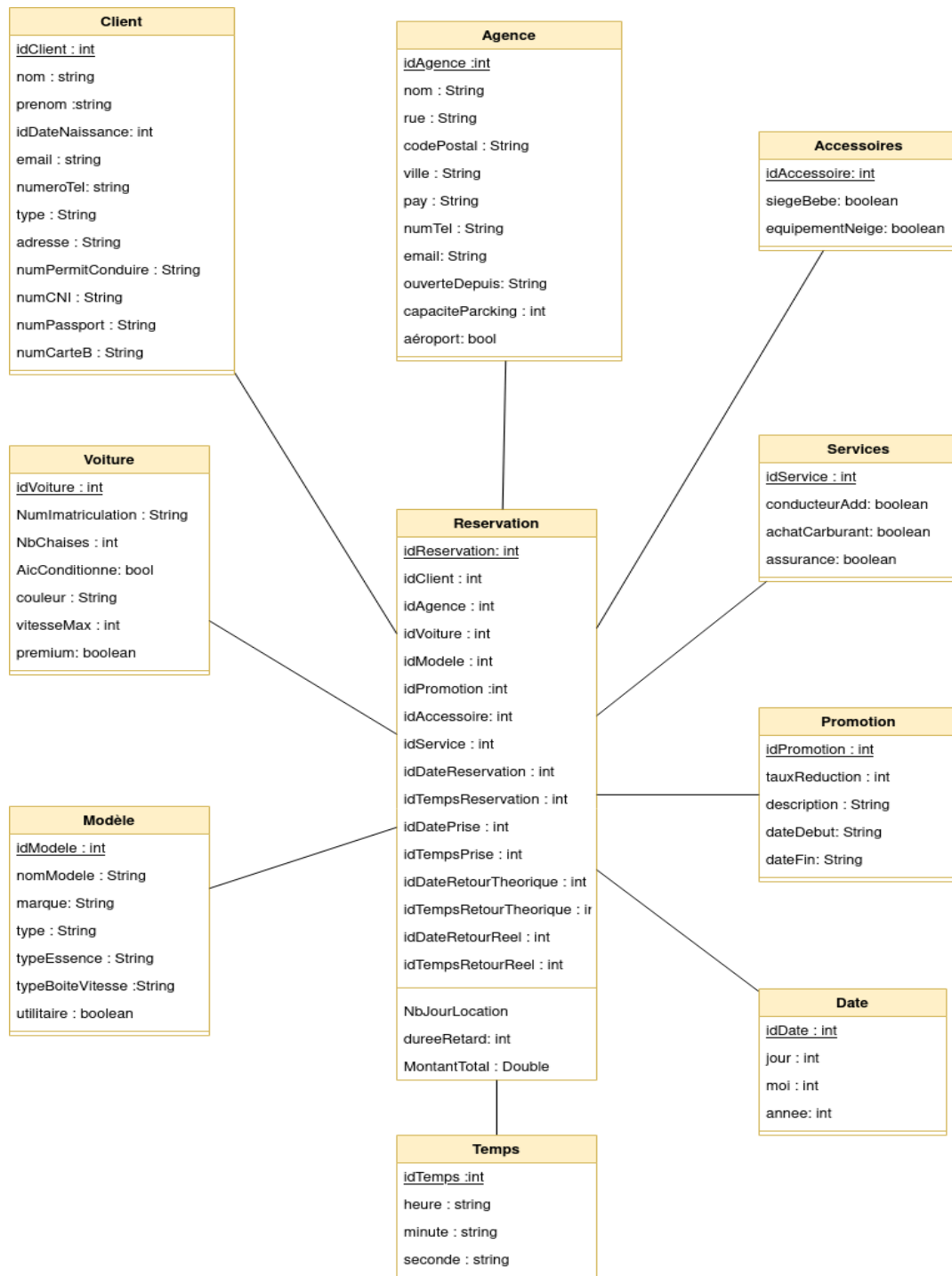
- Un tuple dans la table de faits correspond à une location d'une voiture V de modèle M par un client C dans la date D dans le temps T de la promotion P dans une agence A, payée avec un montant (seulement location) MTL et inclut un ensemble d'accessoires AC et un ensemble de services SE et payé avec un montant total MTT.

La réservation débute dans une date DDebut (avec temps TDebut) et se termine dans une date DFinTheo (avec temps TFinTheo) et DFinReel (avec TFinReel).

Le montant de location MTL, le montant total MTT, le nombre de jours de location NJL et la durée de retard DR correspondent à la mesure de cette table de faits.

Les attributs : client ID, agence ID, date de réservation ID, temps de réservation ID, voiture réservée ID, modèle réservée ID, date de prise en main ID, temps de prise en main ID, date de retour théorique ID, temps de retour théorique ID, date de retour réelle ID, temps de retour réelle ID, promotion ID, accessoires ID, services ID, montant location, montant total, nombre de jours de location, durée de retard.

Modèle logique de l'entrepôt pour le DataMart "RÉSERVATION" :



**b. Pour la deuxième action, prévoyez une modélisation moins détaillée (snapshot, transactions, record-update)**

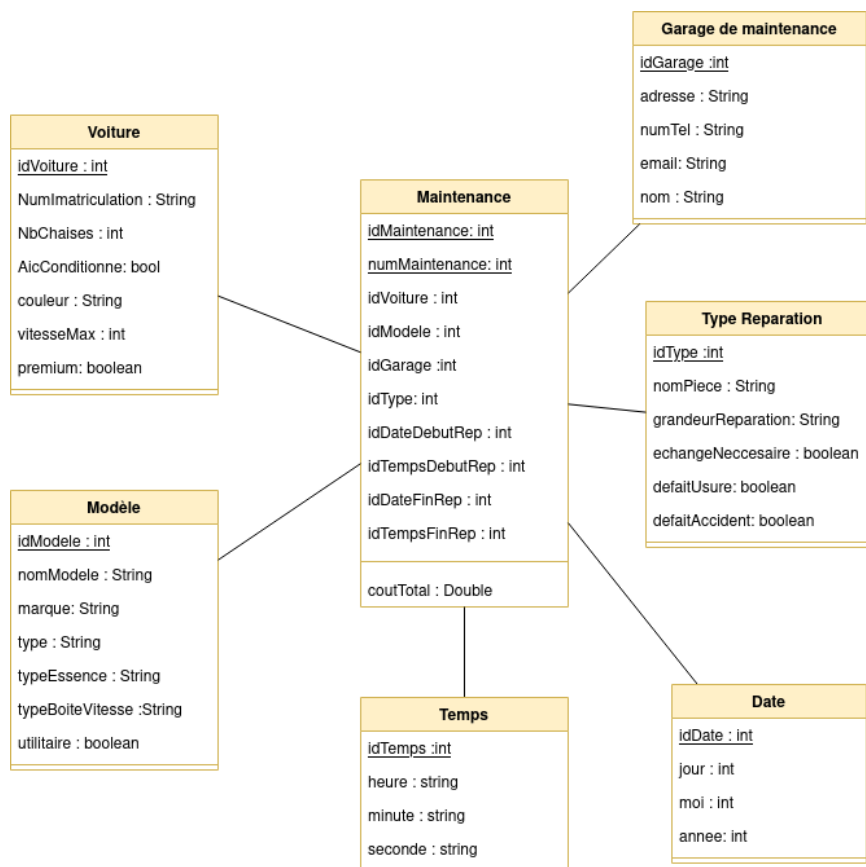
**La maintenance des voitures: On utilisant le modèle RECORD UPDATE**

Un tuple dans la table de faits correspond à une maintenance avec un numéro N d'une voiture V d'un modèle M débute dans une date DDebut (avec temps TDebut) et se termine dans une date

DFin (avec temps TFin) dans un garage de réparation G avec un type de réparation T. La mesure dans ce cas correspond au coût de la maintenance :

numéro, voiture ID, modèle ID, garage ID, typeReparation ID, dateDebutReparationID, dateFinReparationID, tempsDebutReparationID, tempsFinReparationID, coût (mesure)

Modèle logique de l'entrepôt pour le DataMart "Maintenance" :



**c. Pour chaque action / opération, donnez la liste des faits (ou mesures) à enregistrer dans l'entrepôt.**

L'achat des voitures:

Quantité: quantité des voitures achetés par modèle

Montant Total = prix unitaire \* quantité

Distribution des voitures (parking):

Quantité: quantité des voitures par modèle dans un parking d'une agence

**d. Dites s'il s'agit d'une mesure additive, semi-additive, ou nonadditive.**



La réservation / le paiement / la prise en main de voiture: additive, la mesure du montant peut être sommée à travers chaque dimension. En plus, les mesures du nombre de jours de location et la durée de retard peuvent être sommées à travers chaque dimension.

La maintenance des voitures: additive, la mesure du coût peut être sommée à travers chaque dimension

L'achat des voitures: les deux mesures montant total et quantité sont additives

Distribution des voitures (parking): semi-additive.

## **7. Définissez les dimensions nécessaires aux modèles. Pour chaque action, indiquez au moins 5 dimensions.**

Réservation: agence, voiture, modèle, client, date, temps, promotion, accessoires, services

Maintenance: voiture, date, modèle, garage, agence, type, temps

Achat: agence, modèle, date, temps

Distribution: modèle, date, temps, agence

### **a. Pour chaque dimension, prévoir une dizaine d'attributs.**

(...) → remplit plus qu'une colonne

Client: ID, nom, prénom, adresse exacte (...), DateNaissance, email, numeroTel, nationalité, permis de conduire données (...), passeport données (...), carte de crédit données (...), type de client

voiture: ID, plaque d'immatriculation, couleur, , id\_modele, nombre de chaises, premium?, air conditionné?, vitesse Max

agence: ID, nom, rue, code postale, ville, région, pays, numéro tél, email, ouvert depuis, nombre de parking, aéroport?,

date: ID, date, jour de semaine, mois d'année, année, holiday?, weekday?, last day of a week?, last day of a month?, jour populaire?

temps: ID, temps, heure, minutes, secondes, AM/PM, part de journée, temps populaires?

Garage: ID, adresse\_exacte (...), numero\_Tel, adresse e-mail, nom

Type: ID, grandeur de la réparation, nom pièce, échange nécessaire, défaut d'usure, défaut d'accident

Modèle: ID, marque, utilitaire?, type (électrique, hybride, normale) , type d'essence, type de boîte de vitesse (manuelle, automatique)

promotion: ID, taux de reduction, description, date debut, date fin

accessoires: ID, porte bébé, siège enfant, rehausseur, pneus d'hiver, chaines, chaussettes

services: ID, conducteur additionnel, achat de carburant, supercover, personnel assurance, brise glace/crevaison, assistance premium

**8. Est-il possible de répondre aux traitements que vous avez indiqué avec le modèle que vous avez mis en place ? Expliquez pourquoi et comment.**

#### **reservation**

Le nombre de réservations groupées par lieux pour une période définie: oui, le fait table « réservation » à des dimensions date et agence.

Select count (reservation Id) , nom\_agence, from Réservation, Agence, Date

Where Date\_ID < x and Date\_ID > y and Agence.ID = Réservation.agence\_id and Réservation.date\_id = Date.ID

Group by Agence;

Le nombre de paiements effectués par type de client (degré de fidélité) pour une période donnée: Oui, fait table « réservation » à une dimension client .

Select count (reservation Id) , typeClient from Réservation, Client WHERE

Client.ID = Réservation.client\_id and Client.typeClient = x

Group by Client;

Le montant total des paiements par modèle pour une période définie: Oui, la voiture réservée et le montant sont dans le fait table de la réservation et on trouve le modèle dans la dimension voiture.

Select sum(montant) , nom\_modele

From Reservation, Voiture

Where Reservation.voiture\_reservee\_id = Voiture.ID AND Modele.nom\_modele=x

Group by MODELE

#### **Distribution**

Nombre total des voitures stockées par agence: Oui par la dimension « agence » et on peut afficher la mesure quantité de la table de fait Distribution.

Select Quantité

From Distribution, Agence

Where Distribution.agence\_id = agence\_id

Group by Agence

Nombre total des voitures stockées dans un pays particulier: Oui par la dimension « agence » on peut séparer par pays et on peut afficher la mesure quantité de la table de fait Distribution.

Select sum(Quantité), pays

From Distribution, Agence

Where Distribution.agence\_id = agence\_id and Agence.agence\_pays = x

Group by Agence

Les agences qui stockent le plus grand nombre des voitures:

Select sum(Quantité)

From Distribution, Agence

Where Distribution.agence\_id = agence\_id

Group by Agence

Order by count(voiture\_id) Desc Limit 5;

## **Achat des voiture**

Montant d'achat de voiture d'une année

Select sum(montant) from Achat, Date

Where Date\_ID < x and Date\_ID > y and Achat.date\_id = Date.ID

Montant d'achat de voiture d'une année groupé par modèle

Select sum(montant),modèle from Achat, Date

Where Date\_ID < x and Date\_ID > y and Achat.date\_id = Date.ID

Group by idModele

Quantités des voitures électriques achetées par an: oui on peut le répondre, parce que le type du véhicule est notée dans la dimension « voiture ».

```
SELECT sum(Quantité)
```

```
From Achat, Voiture, Date
```

```
Where Date_ID < x and Date_ID > y and Achat.date_id = Date.ID and Achat.voiture_id =  
Voiture.ID and Voiture.voiture_electrique = true
```

```
Group by voiture;
```

### **la maintenance**

La somme des dépenses de maintenance par modèle de véhicule pour une année:

```
Select sum(cout), modèle from Maintenance, Date
```

```
Where Date_ID < x and Date_ID > y and Achat.date_id = Date.ID
```

```
Group by modele
```

La somme des dépenses de maintenance par agence pour une période donnée:

```
Select sum(coût), agence_ID
```

```
From Maintenance, Agence, Date
```

```
Where Maintenance.agence_id= Agence.ID and Maintenance.date_id = Date.ID and date < x and  
date > y
```

```
Group by agence_ID;
```

Montant total des maintenances effectués pour une année:

```
Select sum(coût), année
```

```
From Maintenance, Date
```

```
Where Maintenance.date_id = Date.ID
```

```
Group by année ;
```

**9. Pour tester la pertinence de votre modèle, donner un exemple d'instance de l'entrepôt de données (2 ou 3 lignes par table suffisent).**

```
INSERT INTO Client (idClient, nom, prenom, adresse_exacte, DateNaissance, email, numeroTel,
nationalite, numPermitConduire, numCNI, numCarteB, types)
VALUES
```

```
(5, 'Meflah', 'wided', 'cite des roses', '20/03/2003', 'wided@etu.umontpellier.fr', '003365414141',
'française', '25362255226654', '524161516116', '253654552566355544', 'normale' ),
(65, 'Marie', 'Lon', 'cite des palmes', '10/06/2000', 'marie@etu.umontpellier.fr', '003375475771',
'française', '25362255226654', '524161516116', '253654552566355544', 'normale'),
(99, 'Mark', 'Luk', 'montpellier', '13/12/1999', 'wided@etu.umontpellier.fr', '003368757545',
'française', '25362255226654', '524161516116', '253654552566355544', 'normale' );
```

```
INSERT INTO Modele (idModele, marque, ligne, utilitaire, types, type_d_essence,
ttype_boîte_vitesse) VALUES
```

```
(1, 'renault', 'Kangoo', false, 'électrique', '-', 'manuelle'),
(2, 'audi', 'A5', true, 'hybride', 'gazole', 'manuelle'),
(3, 'peugeot', 'traveller', false, 'normale', 'tout', 'auto');
```

```
INSERT INTO Voiture (idVoiture, Numimatriculation, couleur, id_modele, Nbchaises, premium,
AirConditionne, vitesseMax)
```

```
VALUES
(22, "32 ABC 09", 'rouge', 339, 2, false, false, 250),
(33, "55 XYZ 65", 'bleu', 506, 5, true, false, 190);
```

```
INSERT INTO Agence (idAgence, nom, rue, code postale, ville, région, pays, ouvertDepuis,
capaciteParking, aéroport?,)
```

```
VALUES
(66, 'agence occitanie', 'rue de liberté', '34000', 'montpellier', 'occitanie', 'france', '24/12/2009',
'2', 'false'),
(75, 'agence saint', 'route de l'université', '75000', 'paris', 'île-de-France', 'france', '24/12/2012',
'3', 'false');
```

```
INSERT INTO Garage (idGarage, adresse, numTel , email, nom, capacité)
```

```
VALUES
(2, 'rue rebuffy', '046802571', 'garage.sainteloi@gmail.com', 'garage saint-eloi', 128 ),
(1, 'rue de liberté', '046833371', 'garage.saintdennis@gmail.com', 'garage saint-denis', 30);
```

```
INSERT INTO Dates (idDate, dates, jourSemaine, jour, mois, année, holiday, weekDay,
lastDayofWeek, lastDayMonth, jourPopulaire)
```

```
VALUES
(22052021, '22/05/2021', 'jeudi', 22, 'mai', 2021, 'false', 'false', 'false', 'false', 'false'),
```

(12082021,'12/08/2020','vendredi', 12,'août', 2020,'false', 'false', 'false', 'false','false');

INSERT INTO temps (idTemps, heure, minute, seconde, AMPM, partJournée, tempsPopulaires)  
VALUES

(1, 10, 35, 59, AM, matin, true),  
(2, 22, 08, 37, PM, nuit, false);

INSERT INTO Reservation (idReservation, idClient, idAgence, idVoiture, idModele, idPromotion, idAccessoire, idService, idDateReservation, idTempsReservation, idDatePrise, idTempsPrise, idDateRetourTheorique, idTempsRetourTheorique, idDateRetourReel, idTempsRetourReel, montantLocation, montantTotal, NbJourLocation, idDureeRetard, modePaiement, annulé)  
VALUES

(23, 65,75,25,9995,1,1,256, 5115,09, 65, 339, 78, 65, 007, 115, "245.99", "268.01" 1, "00:00:00", "credit card", false)  
(12, 5,33,66,5254,1,1,256, 427,45, 10, 11, 12, 13, 14, 15, "59.00", "59.00", 6, "23:05:48", "espèces", false),  
(55, 99,33,75,544,1,55,54, 447, 447, 58, 269, 004, 459, 034, 78, "69.99", 69.99" 0, 132, "credit card",true);

INSERT INTO Maintenance (IdMaintenance, numMaintenance, idVoiture, idModele, idGarage, idType, idDateDebutRep, idTempsDebutRep, idDateFinRep, idTempsFinRep, coutTotal)  
VALUES

(23,65,75,25,454,4,6,777, 2757, 4541, "999.58"),  
(5,17,33,66,754,1,1,754, 5745, 4545, "45.06"),  
(99,40, 33,75,57,1,57,744, 37,4545, "1009.89");

INSERT INTO Type: (idType, grandeurRéparation, nomPièce, échangeNécessaire, défautUsure, défautAccident)  
VALUES

(1, "petite", "allumeur", true, true, false),  
(2, "grande", "roue 3", true, false, true);

INSERT INTO promotion: (idPromotion, tauxReduction, description, dateDebut, dateFin)  
VALUES

(1, "15.5%", "black friday reduction 21", '01/09/2021', '01/11/2021'),

(2, "30.0%", "valentines day", '13/02/22', '15/02/2022');

INSERT INTO accessoire: (idAccessoire, siègeBébé, siègeEnfant, rehausseur, équipementNeige)  
VALUES

(1, false, false, false, false),

(2, true, false, false, false);

INSERT INTO service: (idService, conducteurAdditionnel, achatCarburant, assurance)  
VALUES

(1, false, false, 0),

(2, false, false, 1);

**10. Estimez la taille des tables de l'entrepôt (en terme du nombre de lignes) sur 12 mois. Est ce que cette taille est raisonnable ? Justifiez votre réponse**

Réservation:

Hertz avait 2,5 Mrd \$ de revenus dans les derniers trois mois globalement

→environ 10 Mrd \$ de revenus pour l'année

On estime la moyenne coût de location à 600\$

→environ 17 Mio de locations pour l'année

→environ 17 Mio de lignes pour l'année

Cette taille est raisonnable, comme Réservation est une table de faits et donc construite pour avoir beaucoup de lignes.

Agence:

Hertz a environ 10.000 agences globalement

→environ 10.000 de lignes

Cette taille est raisonnable, comme elle n'est pas trop grande pour une dimension ( $\ll 10^6$ ).

Voiture:

Hertz possède environ 424.000 voitures dans les Etats-Unis

On estime que l'entreprise a entre 1,5 Mio voitures en tout

→ 1,5 Mio lignes

Cette taille est problématique, comme les dimensions devraient avoir moins d'une Mio de lignes.

D'autre côté, les voitures sont une des dimensions essentielles pour Hertz ce qui peut rendre

raisonnable une telle taille. On pourrait manier cette dimension par exemple via “partitioning by row”, où on divise les voitures en plusieurs dimensions pour de différents types de voiture.

Date:

Une année a au plus 366 journées

→366 lignes

Cette taille est très petite et ne pose aucun problème pour la table de Date.

Temps:

60 secondes \* 60 minutes \* 24 heures = 86.400 secondes

→86.400 lignes

Cette taille est raisonnable, comme elle n’est pas trop grande pour une dimension ( $\ll 10^6$ ).

Garage:

On estime qu’il y a environ 3 garages pour chaque agence

10.000 agences \* 3 garages = 30.000 garages

→30.000 lignes

Cette taille est raisonnable, comme elle n’est pas trop grande pour une dimension ( $\ll 10^6$ ).

Type:

Une voiture a environ 1.800 (grandes) parts différentes

Si on a besoin de chaque combinaison de “grandeur de la réparation”, “échange nécessaire” et “défait d’usure ou défaut d’accident” on reçoit

$1.800 * 2 * 2 * 2 = 14.400$  lignes

Cette taille est raisonnable, comme elle n’est pas trop grande pour une dimension ( $\ll 10^6$ ).

Promotions:

On estime que Hertz fait environ 6 promotions différentes chaque année

→6 lignes

Cette taille est raisonnable, comme elle n’est pas trop grande pour une dimension ( $\ll 10^6$ ).

Accessoires:

On a besoin d’au plus toutes les combinaisons (0 ou 1) de “porte bébé”, “siège enfant”, “rehausseur”, “pneus d’hiver”, “chaines” et “chaussettes”

→ $2 * 2 * 2 * 2 * 2 * 2 = 64$  lignes

Cette taille est raisonnable, comme elle n’est pas trop grande pour une dimension ( $\ll 10^6$ ).

Services:

On a besoin d’au plus toutes les combinaisons (0 ou 1) de “conducteur additionnel”, “achat de carburant”, “assurance”

→ $2 * 2 * 2 = 8$  lignes

Cette taille est raisonnable, comme elle n’est pas trop grande pour une dimension ( $\ll 10^6$ ).



### Clients:

On estime qu'un client fait 4 locations par an en moyenne

$16.700.000 \text{ locations} / 4 \text{ locations par client} = 4.175.000 \text{ clients}$

→ Environ 4 Mio lignes

Cette taille est problématique, comme les dimensions devraient avoir moins d'une Mio de lignes. Les clients sont souvent une très grande et très détaillée dimension, comme elle est une des plus importantes. On pourrait manier cette dimension par exemple via "partitioning by column", où on divise les attributs des clients en plusieurs dimensions pour de différents degrés de variabilité.

## **12. Proposez 10 requêtes analytiques correspondant aux traitements que vous avez indiqués.**

### Q1 : Le nombre de réservations par agence et pay de chaque année

Select count (reservationId) , nom\_agence, pay, Date.year from Réservation, Agence, Date

Where Agence.ID = Réservation.agence\_id and Réservation.dateReservation\_id = Date.ID

Group by AgenceId, pay, Date.year;

### Q2 : Le montant total des paiements effectués par type de client (degré de fidélité) par agence par moi

Select sum (montantTotal) , typeClient, Date.year from Réservation, Client WHERE

Client.ID = Réservation.client\_id

Group by Client.type, agenceId, Date.year;

### Q3:Le nombre des réservations par modèle par agence par mois par année

Select count(reservationId) , nom\_modele, nomAgence

From Reservation, modele, Agence, Date

Where Reservation.modelId= modele.ID AND agenceId= Agence.id

AND idDateReservation = Date.id

Group by idModele, idAgence, Date.month, Date.year

Q4 : Nombre total des voitures stockées par modèle, agence par pays

```
Select sum(Quantité), Modèle.nom, Agence.nom, pays From Distribution, Agence
Where Distribution.agenceId = Agence.Id and Distribution.modeleId = Modele.id
Group by Modele.id, Agence.id, pay
```

Q5 : Les agences qui stockent le plus grand nombre des voitures:

```
Select sum(Quantité)
From Distribution, Agence
Where Distribution.agenceId = agenceId
Group by Agence
Order by count(agenceId) Desc Limit 50;
```

Q6 : Montant total des achats de voitures groupé par modèle par année

```
Select sum(montant),modèle from Achat, Date
Where achat.modele_id=Modele.id and Achat.date_id = Date.ID
Group by Modele.id, Date.year
```

Q7 : Quantités des voitures électriques achetées par année:

```
SELECT sum(Quantité),année
From Achat, Voiture, Date
Where Achat.date_id = Date.ID and Achat.voiture_id = Voiture.ID and
Voiture.voiture_electrique = true
Group by voiture_id,Date.year
```

Q8 : La somme des dépenses de maintenance par modèle de véhicule par année:

Select sum(cout), modèle, année from Maintenance, Date

Achat.date\_id = Date.ID and Achat.modele\_id=Modele.ID

Group by modele\_id, Date.année

Q9-La somme des dépenses de maintenance par agence par année:

Select sum(coût), agence.nom, Date.année

From Maintenance, Agence, Date

Where Maintenance.agence\_id= Agence.ID and Maintenance.date\_id = Date.ID

Group by agence\_id, Date.année

Q10-Montant total des maintenances effectués par année:

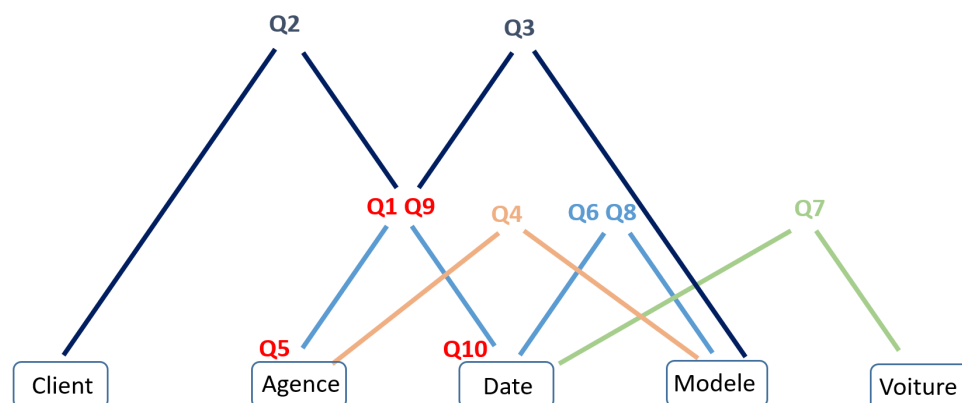
Select sum(coût), année

From Maintenance, Date

Where Maintenance.date\_id = Date.ID

Group by Date.année ;

**13. Donnez l'ensemble des vues matérialisées permettant de répondre à l'ensemble de vos requêtes. Utiliser la technique basée sur le treillis d'agrégation présentée en cours.**



sources:

<https://thepointsguy.com/news/hertz-status-promotion/>

<https://www.hertz.com/rentacar/productsandservices/productsandservicesRegions.do>

<https://moneyinc.com/biggest-rental-car-companies-in-the-world/>

[https://en.wikipedia.org/wiki/The\\_Hertz\\_Corporation](https://en.wikipedia.org/wiki/The_Hertz_Corporation)

<https://www.asyousow.org/resolutions/2019/12/12/hertz-climate-change-risk-reporting>