



ENSA  
ÉCOLE NATIONALE DES SCIENCES  
APPLIQUÉES  
KHOURIBGA



# RAPPORT

Module : Système d'aide à la décision  
1<sup>ère</sup> année Master Big Data et Aide à la Décision

---

## Modélisation d'un système décisionnel pour une compagnie d'assurance de biens.

---

**Réalisé par :**

CHANAA Ousama  
BELAKTIRIA Khadija

**Encadré par :**

Mme SOUSSI Nassima

# Table des matières

Introduction .....	4
Description des besoins .....	5
Les acteurs : .....	<b>5</b>
Les besoins fonctionnels : .....	<b>5</b>
Système transactionnel : .....	5
Entrepôt de données : .....	6
Conclusion : .....	<b>6</b>
Conception.....	7
Modèle relationnel de données : .....	<b>7</b>
Table client : .....	8
Table opérateur : .....	8
Table type_opérateur : .....	8
Table tiere_impliques : .....	8
Table expertise : .....	9
Table risque : .....	9
Table note : .....	9
Table bien : .....	9
Table paiement : .....	10
Table Police : .....	10
Table état : .....	10
Table sinistre : .....	10
Table transaction : .....	11
Table date : .....	11
Conclusion : .....	<b>11</b>
Alimentation de l'entrepôt de données : .....	12
Conception de l'entrepôt de données : .....	<b>12</b>
Tables de faits : .....	12
Mesures : .....	12
Dimensions : .....	12
Schéma en Constellation : .....	15
Processus ETL : .....	<b>15</b>
Extraction des données : .....	16
Transformation des données : .....	17
Chargement des données : .....	17

Conclusion :.....	<b>17</b>
Schéma Mondrian.....	18
Schéma Mondrian :.....	<b>18</b>
Restitution des vues avec le Pivot4j : .....	<b>19</b>
Exemples de vues : .....	19
Conclusion :.....	<b>21</b>
Implémentation des tableaux de bord.....	22
Power BI :.....	<b>22</b>
Tableau de bord.....	<b>22</b>
Conclusion générale .....	<b>24</b>

# Introduction

Les systèmes d'aide à la décision sont des programmes informatisés utilisés pour appuyer les déterminations, les jugements et les lignes de conduite des organisations et des entreprises.

Les DSS passent au crible et analysent des quantités massives de données, compilant des informations complètes qui peuvent être utilisées pour résoudre des problèmes et prendre des décisions.

En s'appuyant sur une grande quantité de données, des historiques de toutes transactions réalisées, un système d'aide à la décision produit des rapports d'informations complets. Ainsi qu'il permet aux utilisateurs de prendre des décisions plus éclairées, plus efficaces et plus rapides.

Un système d'aide à la décision, est utilisé par la direction des opérations et par les départements de planification pour compiler des informations et des données et les synthétiser en enregistrement.

Notamment, Une compagnie d'assurance peut l'utiliser pour concevoir les transactions effectuées par ses employés, ainsi que toutes les estimations possibles.

Afin de faciliter la prise de décisions, les données sont représentées sous formes claires, simples et représentatives. Par exemple les tableaux de bord, qui peuvent être analysés suivant différents composants tel que par client, par police, par sinistre, par bien...

Dans ce rapport nous implémenterons un système d'aide à la décision qui envisage la représentation des données sous formes de tableaux de bord, en passant par la conception d'une base de données source, le passage au datawarehouse, et la représentation des données sous forme de vues analysables. En terminant par la construction des tableaux de bord voulus.

## Description des besoins

Les compagnies d'assurances sont des entreprises qui fournissent des services d'assurance de biens à des clients qui deviennent des assurés.

Les biens peuvent être de différents types, tels que les immobiliers, les automobiles ou les responsabilités civiles.

Donc ces compagnies doivent gérer les contrats des clients ainsi que les accidents que déclarent ces clients. D'où l'obligation de se disposer d'une application transactionnelle permettant la bonne gestion des tâches déclarées récemment.

### Les acteurs :

Un acteur est une entité utilisateur externe (utilisateur humain ou autres système) qui interagit directement avec le système étudié. En réponse à l'action d'un acteur, le système fournit un service qui correspond à son besoin :

- **Opérateur** : peut être employé de l'assurance ou un agent externe travaillant avec la compagnie, c'est l'acteur responsable de la manipulation des transactions
- **Client** : c'est la personne qui veut avoir une police avec la compagnie, afin d'assurer un ou plusieurs biens.
- **Tiers impliqués dans le sinistre** : est la personne avec laquelle le client fait l'accident.
- **Expertise** : l'expert qui évalue les dommages accidentels.

### Les besoins fonctionnels :

Après une étude détaillée du système, cette partie est réservée à la description des exigences fonctionnelles demandées.

Les besoins sont divisés en deux parties :

#### Système transactionnel :

Chaque opérateur peut effectuer des transactions, pour police tout en ayant le droit des trois actions principales. Cela est valide pour une police, un risque et des biens assurés sur un risque. L'opérateur peut aussi chiffrer ou refuser le risque en question et faire la validation ou le refus de la police. Comme il peut faire des transactions pour les sinistres tout en effectuant les 3 actions pour une expertise, les paiement le sinistre.

La transaction est définie par son identifiant, sa date, une date de début de l'assurance, le client, l'opérateur, le risque voulu, le bien assuré et le numéro de la police en question, le tiers impliqué dans le sinistre, le montant financier.

Le client peut être une personne physique qui assure son bien, ou une personne morale (ex : entreprise).

La police est le contrat signé entre le client et la compagnie, et elle couvre un bien pour un risque donné.

Le sinistre est l'accident fait par le client, il concerne une police et un bien donné.

On entend dire de gérer les trois actions principales sur un objet : l'ajout, la suppression, et la modification.

### **Entrepôt de données :**

Cette application transactionnelle servira à créer un entrepôt de données. L'entrepôt de données, doit aussi donner les informations suivantes :

- Le montant de la prime pour chaque bien assurer, ainsi que le nombre de paiements.
- L'état de la police ainsi que le sinistre en relation avec.

Les informations extraites sont représentées sous forme de tableaux de bord, par client, opérateur, date d'effet, police et par risque et des tableaux de bord de l'activité sur le dossier, notamment, sur le nombre de transactions ou de sinistres.

### **Conclusion :**

Dans ce chapitre, nous avons fait une étude et analyse des besoins. Nous avons défini les différents besoins fonctionnels. Le chapitre suivant sera consacré pour la conception de notre projet.

# Conception

Dans ce chapitre, nous abordons la partie conception du projet, nous détaillons les différents éléments de la conception des bases de données.

## Modèle relationnel de données :

Un modèle relationnel est une structure de données sous forme de tables ou de relations. C'est le modèle le plus répandu dans les SGBD relationnels.

La figure 1 représente notre base de données relationnelle.

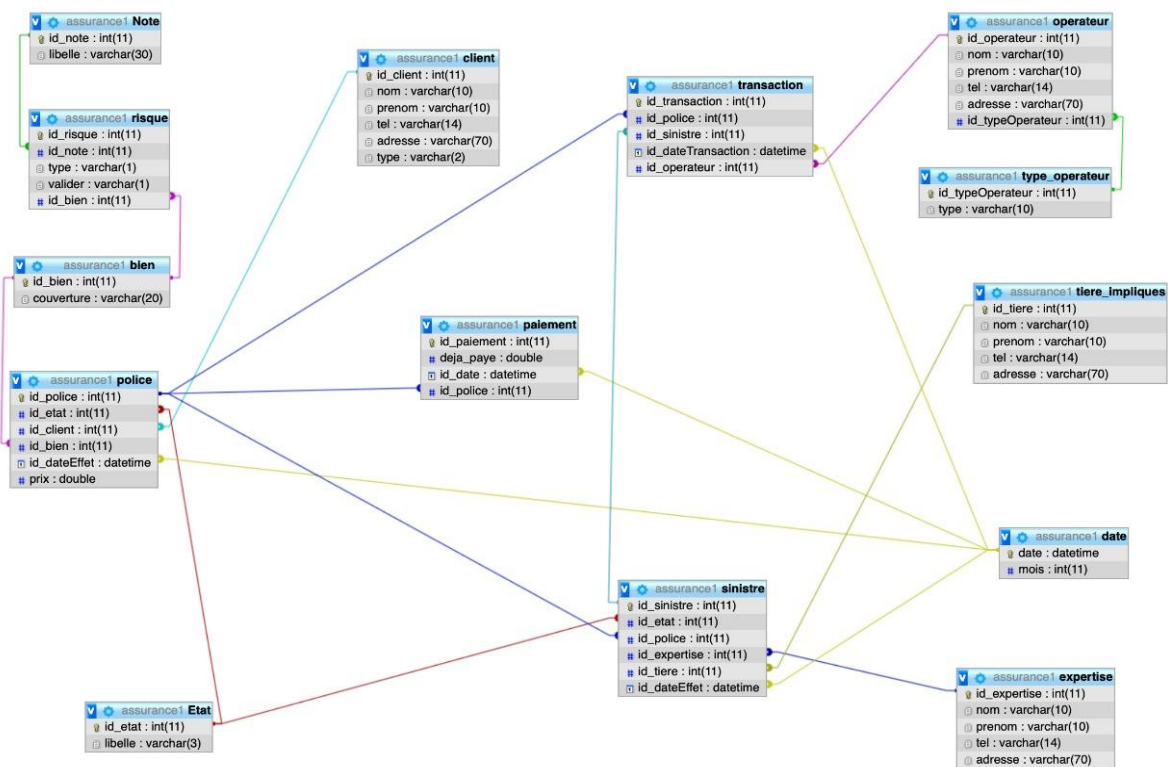


Figure 1 : modèle relationnel de la base de données assurance

Les tableaux suivants représentent les tables utilisées dans la base de données :

Table client :

Attribut	Désignation	Type
<b>id_client</b>	Identifiant du client	Int
<b>Nom</b>	Nom du client	Varchar(10)
<b>Prénom</b>	Prénom du client	Varchar(10)
<b>Tel</b>	Numéro de téléphone avec indicatif	Varchar(14)
<b>Type</b>	Personne physique ou morale	Varchar(2)

Table opérateur :

Attribut	Désignation	Type
<b>id_opérateur</b>	Identifiant de l'opérateur	Int(11)
<b>Nom</b>	Nom de l'opérateur	Varchar(10)
<b>Prénom</b>	Prénom de l'opérateur	Varchar(10)
<b>Tel</b>	Numéro de téléphone avec indicatif	Varchar(14)
<b>Id_typeOperateur</b>	Clé étrangère de la table type_opérateur	int(11)

Table type\_opérateur :

Attribut	Désignation	Type
<b>id_typeOperateur</b>	Identifiant du type	Int(11)
<b>Type</b>	Si l'opérateur est un employé ou un agent externe	Varchar(10)

Table tiers\_impliqués :

Attribut	Désignation	Type
<b>id_tiers</b>	Identifiant du tiers impliqué	Int(11)
<b>Nom</b>	Nom du tiers impliqué	Varchar(10)
<b>Prénom</b>	Prénom du tiers impliqué	Varchar(10)
<b>Tel</b>	Numéro de téléphone avec indicatif	Varchar(14)



Table expertise :

Attribut	Désignation	Type
<b>id_expertise</b>	Identifiant de l'expert	Int(11)
<b>Nom</b>	Nom de l'expert	Varchar(10)
<b>Prénom</b>	Prénom de l'expert	Varchar(10)
<b>Tel</b>	Numéro de téléphone avec indicatif	Varchar(14)

Table risque :

Attribut	Désignation	Type
<b>id_risque</b>	Identifiant du risque	Int(11)
<b>Id_note</b>	Clé étrangère de la table note	Int(11)
<b>Type</b>	Type du risque (voiture, maison)	Varchar(1)
<b>Valider</b>	Si le risque est validé ou non	Varchar(1)

Table note :

Attribut	Désignation	Type
<b>id_note</b>	Identifiant de la note	Int(11)
<b>Libelle</b>	Evalue le degré du risque (si il est élevé, moyen ou faible)	Varchar(30)

Table bien :

Attribut	Désignation	Type
<b>id_bien</b>	Identifiant du bien assuré	Int(11)
<b>Couverture</b>	Type du bien (voiture, maison)	Varchar(20)

Table paiement :

Attribut	Désignation	Type
<b>id_paiement</b>	Identifiant du paiement	Int(11)
<b>Déjà payé</b>	Le montant déjà payé dans la police désignée	Double
<b>Id_date</b>	La date du paiement	Datetime
<b>Id_police</b>	Identifiant de la police désignée	Int(11)

Table Police :

Attribut	Désignation	Type
<b>id_police</b>	Identifiant de la police(contrat)	Int(11)
<b>Id_etat</b>	Signifie si cette police est nouvellement créée ou nouvellement modifiée	int(11)
<b>Id_client</b>	Identifiant du client signant la police	int(11)
<b>Id_bien</b>	Identifiant du bien concerné par cette police	Int(11)
<b>Prix</b>	Prix total à payer dans cette police	Double

Table état :

Attribut	Désignation	Type
<b>id_etat</b>	Identifiant de l'état	Int(11)
<b>libelle</b>	Donne l'information sur la police ou sur le sinistre	Varchar(3)

Table sinistre :

Attribut	Désignation	Type
<b>id_sinistre</b>	Identifiant du sinistre(accident)	Int(11)
<b>Id_etat</b>	Si ce sinistre est clôturé ou encore encours	int(11)
<b>Id_police</b>	Identifiant de la police concernée	int(11)
<b>Id_expertise</b>	L'expert évaluant l'accident	int(11)
<b>Id_dateEffet</b>	Date du sinistre	Datetime

Table transaction :

Attribut	Désignation	Type
<b>Id_transaction</b>	Identifiant de la transaction	Int(11)
<b>Id_police</b>	Identifiant de la police	Int(11)
<b>Id_sinistre</b>	Identifiant du sinistre	Int(11)
<b>Id_dateTransaction</b>	Date quand cette transaction est effectuée	Int(11)

Table date :

Attribut	Désignation	Type
<b>Date</b>	La date	datetime
<b>Mois</b>	Le mois contenu dans la date	int(11)

## Conclusion :

Dans cette partie, nous avons créé la base de données relationnelle de notre projet. Mais les bases de données relationnelles, sont conçues pour être efficaces pour les transactions, et peu structurées pour l'analyse. Ce système est donc peu adapté à la vision à long terme et donc à la prise de décision. D'où l'utilité de transformer la base de données relationnelles assurance en un entrepôt de données.

Pour ce faire, on utilise le processus ETL via Pentaho.

## **Alimentation de l'entrepôt de données :**

L'entrepôt de données est une collection de données orientées sujet, intégrées, non volatiles et organisées, pour un processus d'aide à la décision.

Il est composé de plusieurs Datamarts.

Un data Mart est une forme simple d'un entrepôt de données qui se concentre sur un seul domaine d'activités.

Dans notre projet, on compose deux datamarts, le premier est pour la police et le deuxième est consacré au sinistre.

## **Conception de l'entrepôt de données :**

La conception des datamarts est basée sur la modélisation dimensionnelle dans un schéma composé de mesures, tables de faits et de mesures, le rassemblement des deux datamarts donne en résultat une conception de tout l'entrepôt de données.

### **Tables de faits :**

La table de faits est la table principale pour chaque Data Mart, elle contient les clés étrangères des dimensions, ainsi que les mesures.

On modélise une table sinistre pour le premier Data Mart et une table de fait police pour le deuxième.

### **Mesures :**

Les mesures sont les données analysables selon différents axes.

Pour le sinistre on calcule le nombre de sinistre

Pour la police on calcule le chiffre d'affaires.

### **Dimensions :**

Les dimensions sont les axes suivant lesquels est basée l'analyse des mesures. Elles sont représentées sous formes de tables.

Nous avons constitué donc des dimensions pour la table sinistre :

Opérateur, risque, date, bien et client.

Pour la table de fait police, on définit les dimensions suivantes :

Etat, risque, operateur, client, bien et date.

On construit donc le premier Data Mart représenté sous forme d'un schéma en étoile, qui contient la table de fait sinistre avec les différentes dimensions citées ci-dessous figure 2.

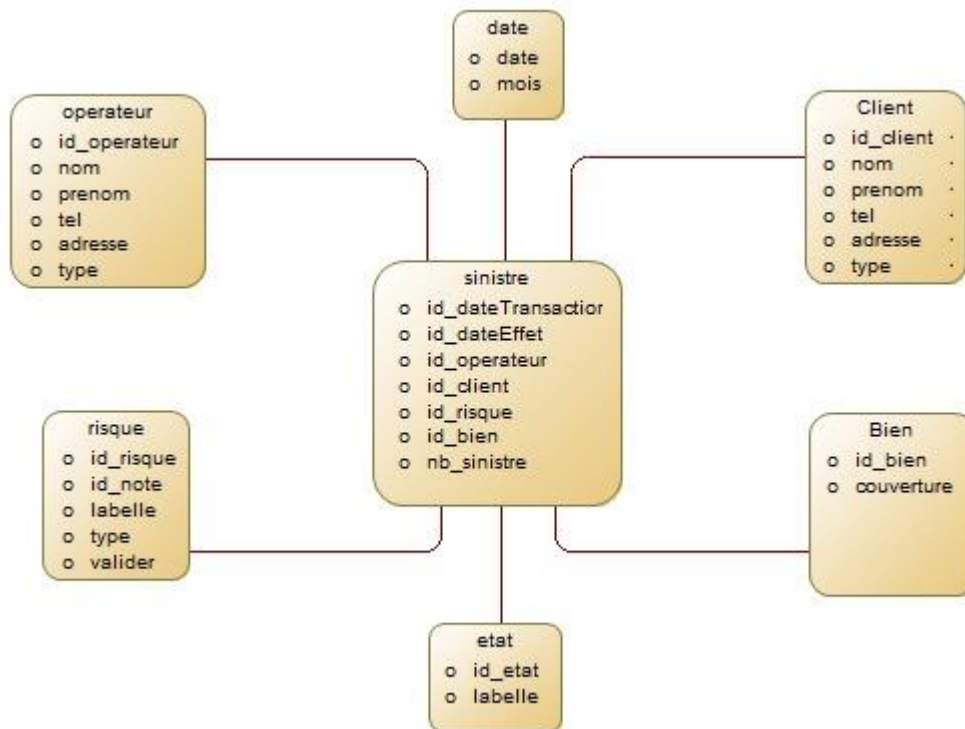


Figure 4 : schéma en étoile de la table de fait sinistre.

On peut décomposer plusieurs dimensions en niveaux hiérarchiques (figure5), ce qui construit un schéma en flocon.

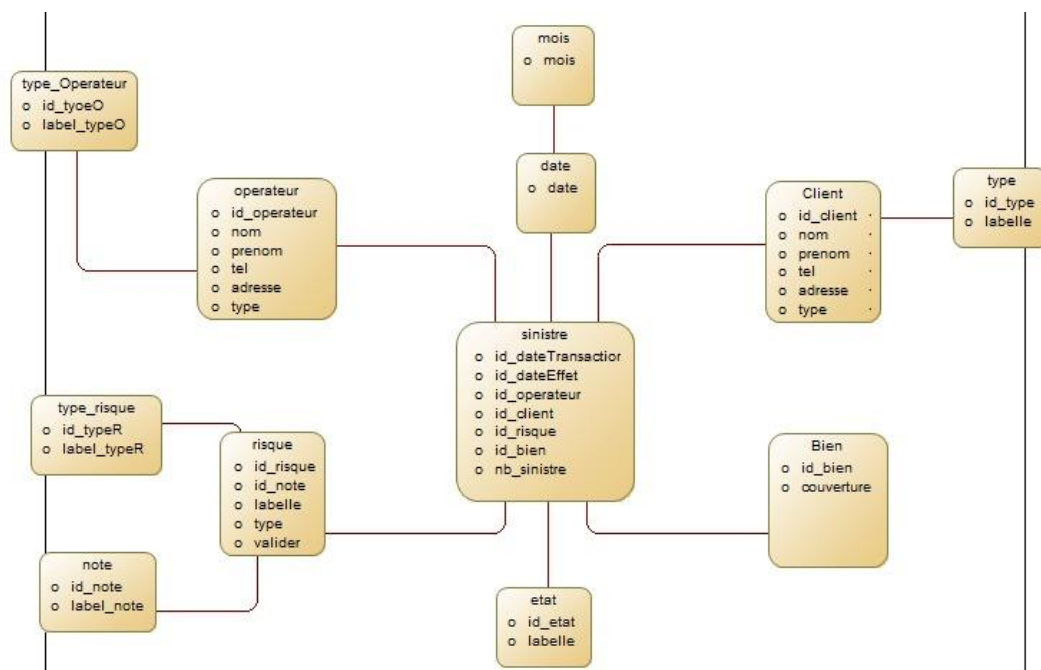


Figure 5 : schéma en flocon de la table sinistre

Le deuxième Data Mart est aussi représenté par un schéma en étoile (figure 6), c'est le schéma de table de fait police, avec les différentes dimensions.

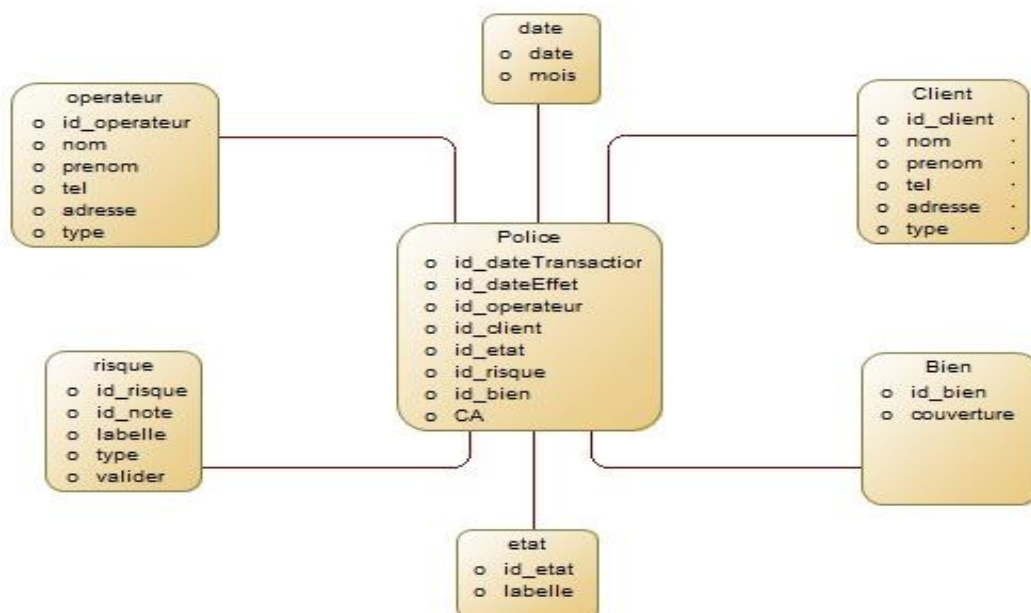


Figure 6 : schéma en étoile de la table police.

Les dimensions liées avec la table de fait police sont, majoritairement, les mêmes suivant lesquels on analyse les sinistres, cela nous permet de construire un schéma en flocon pour la table police aussi (figure 7).

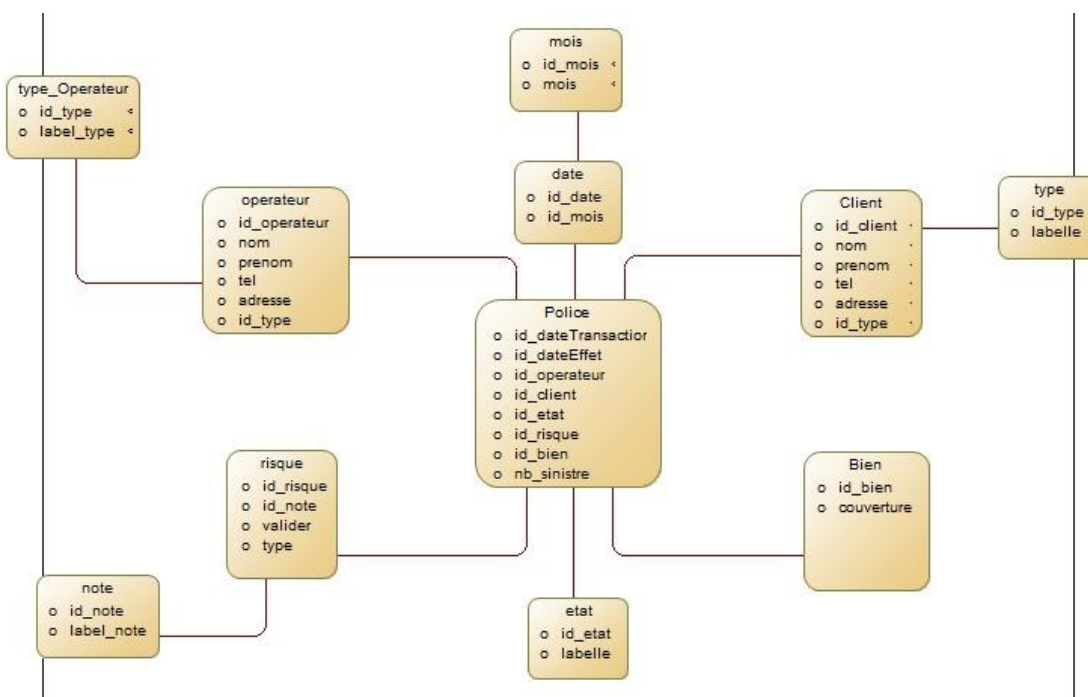


Figure 7 : schéma en flocon pour la table police

## Schéma en Constellation :

Le rassemblement des deux datamarts ou de plusieurs data Mart qui ont des dimensions communes est appelé un schéma en constellation. Cela est équivalent à la conception de notre entrepôt définie dans la figure (8).

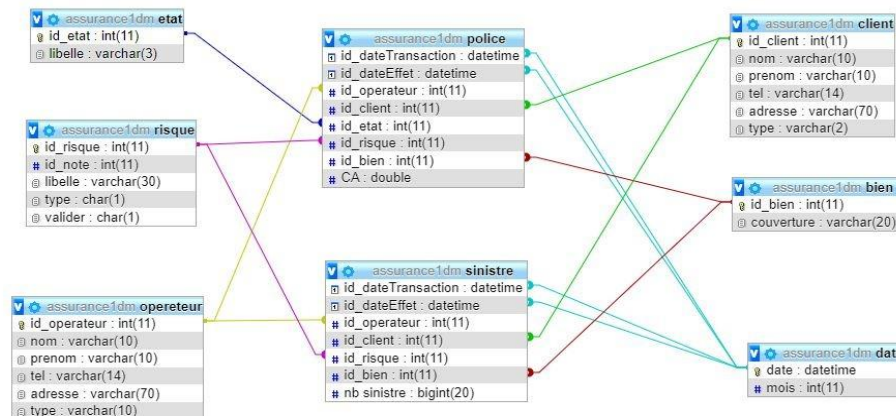


Figure 8 : Conception de l'entrepôt de données en constellation

L'alimentation d'un entreposage des données (datawarehouse) est assurée par le processus ETL.

## Processus ETL :

(Extraction Transformation Loading) qui sert à copier les données depuis le système transactionnel vers le datawarehouse d'une manière automatisée.

Il est composé de trois phases principales :

- Extraction des données
- Nettoyage et transformation
- Chargement

## Extraction des données :

On associe à chaque table de notre base de données relationnelle, jugée utile et judicieuse, un extracteur de données.

La figure (9) et (10) représentent respectivement l'extraction des données depuis la table sinistre et la table police :

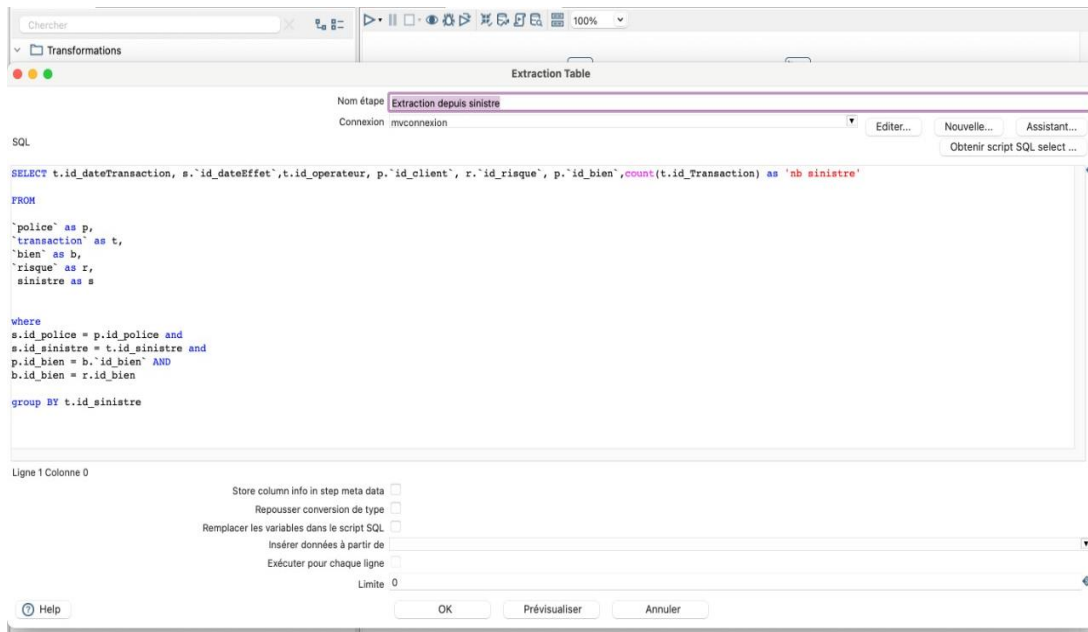


Figure 9 : extraction des données depuis sinistre

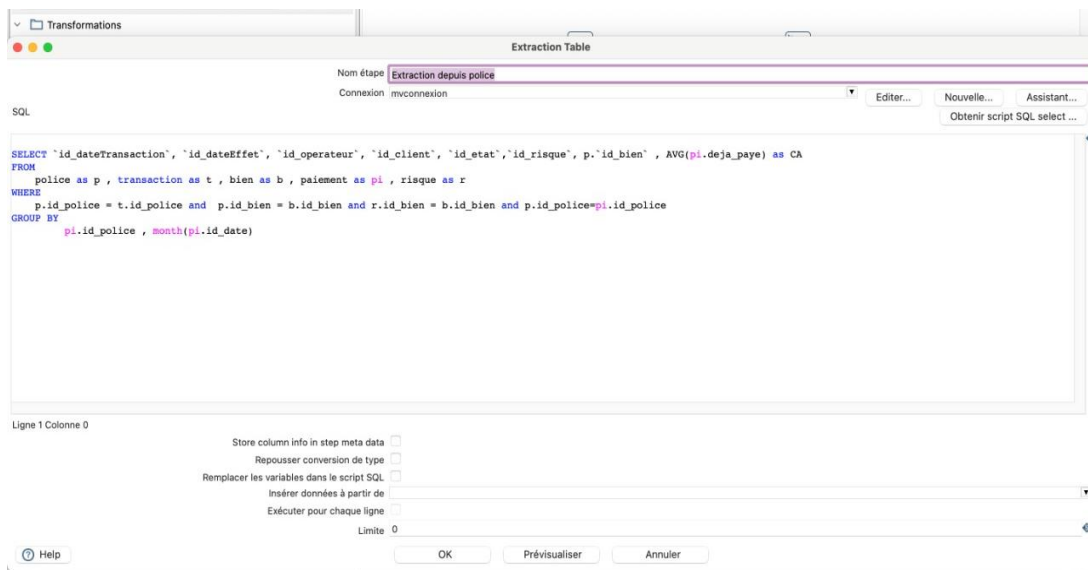


Figure 10 : extraction des données depuis police



## Transformation des données :

Cette étape sert à résoudre les problèmes de consistance des données en éliminant les données fausses, les persistance de données obsolètes...

## Chargement des données :

C'est l'étape permettant de charger les données nettoyées et préparées vers les datamarts.

La figure (11) montre une transformation permettant le chargement de toutes les données dans les datamarts.

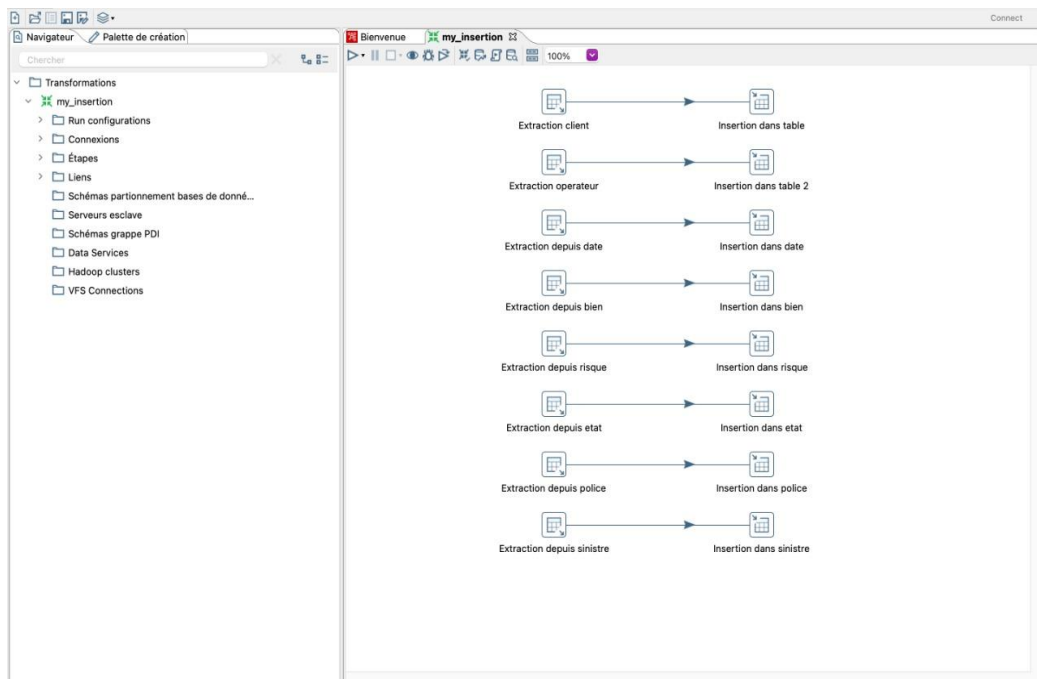


Figure 11 : Chargement des données dans le datawarehouse.

## Conclusion :

Dans cette partie nous avons construit les datamarts, sous forme de schéma en flocons, puis nous l'avons implémenté sous forme d'un schéma en constellation. Ce dernier a été rempli avec les données extraites du système relationnel en faisant appel au processus ETL. Cela donne naissance à une nouvelle partie qui est la création du schéma Mondrian et la publication du cube ROLAP sur le serveur pentaho. Ceci sera l'objet de la partie suivante.

# Schéma Mondrian

OLAP (online Analytical Processing) est une technologie qui donne droit à analyser les données multidimensionnelles, sous forme de cubes ROLAP.

Les cubes ROLAP permettent l'accès direct aux données physiques des tables relationnelles. Cela est implémenté sur un moteur OLAP open source à l'aide d'un fichier XML qui décrit les axes de mesures et les dimensions. C'est la solution Mondrian.

Donc un schéma Mondrian permet de décrire les cubes, les dimensions, les hiérarchies, les niveaux et les membres, ainsi qu'il correspond à la source de données du modèle en étoile ou en flocon.

Ces schémas sont créés et publiés via le concepteur «pentaho schema workbench».

## Schéma Mondrian :

Notre schéma est composé de deux cubes avec les différents niveaux hiérarchiques définis dans les schémas en flocon et en étoile définis ci-dessus.

Les figures 12 et 13 représentent respectivement le schéma créé avec les dimensions et avec les différents niveaux hiérarchiques :

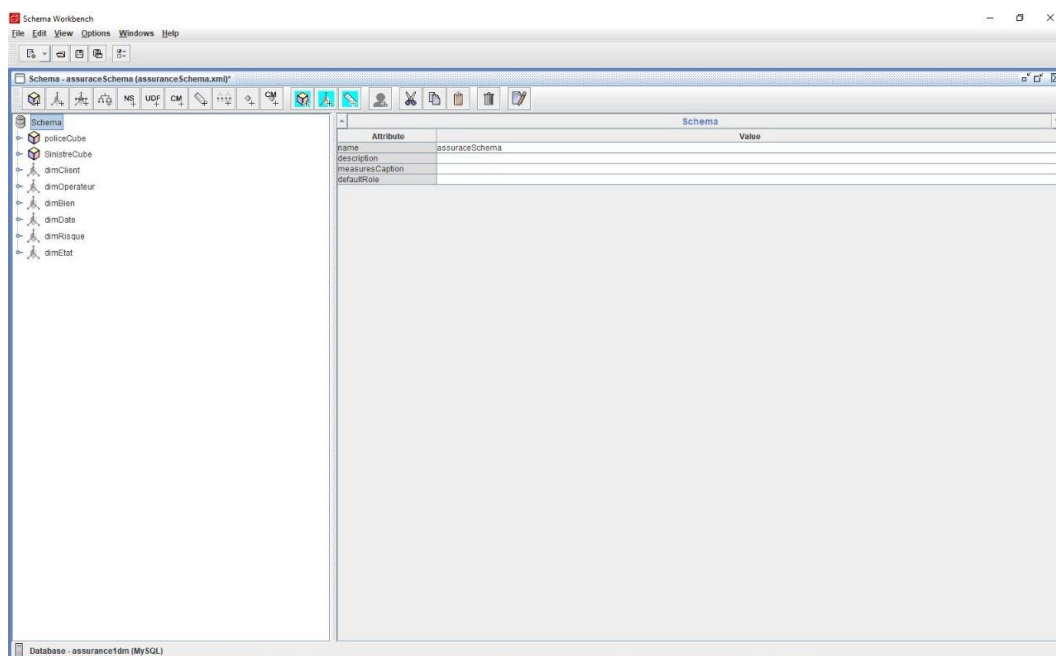


Figure 12 : schéma Mondrian

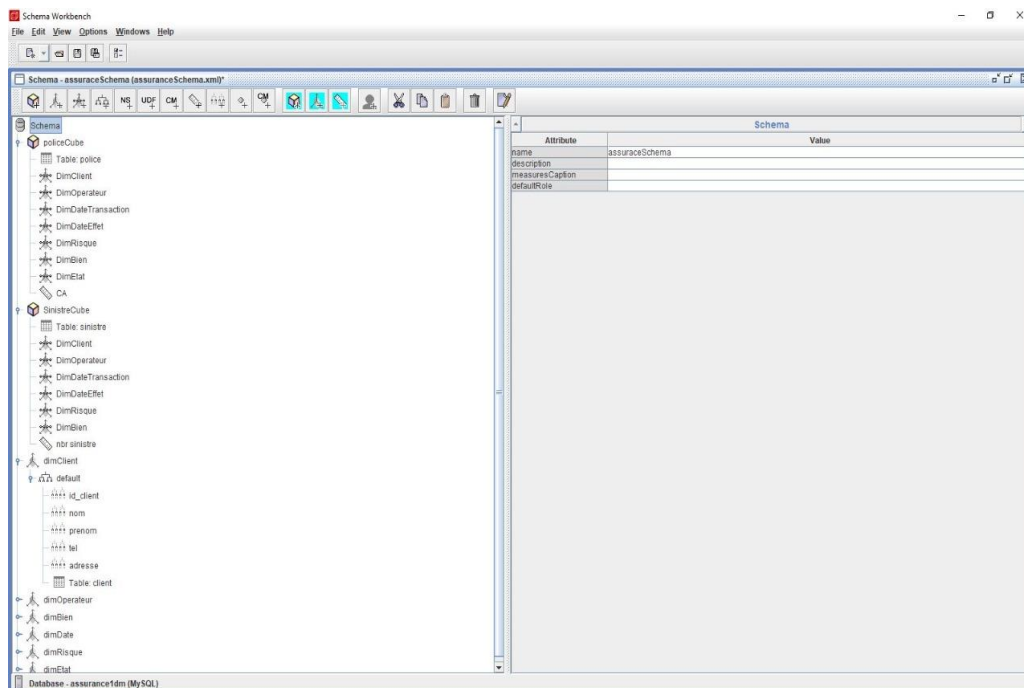


Figure 13 : niveaux hiérarchiques du schéma Mondrian

Après avoir modélisé le schéma à partir de l'entrepôt de données. Une phase de connexion avec le serveur Pentaho est nécessaire pour pouvoir publier notre schéma. L'étape suivante consiste à créer les vues métiers pour exploiter les données. C'est la tâche à effectuer avec le plugin Pivot4j.

## Restitution des vues avec le Pivot4j :

Le pivot4j est un outil de requêtage puissant open source qui permet de représenter le cube sous forme d'un tableau croisé, d'insérer les différentes dimensions afin d'étudier les mesures du cube clairement et facilement.

## Exemples de vues :

On calcule le nombre de biens pour chaque client, ainsi que le total que ce client a payé annuellement (Figure14)

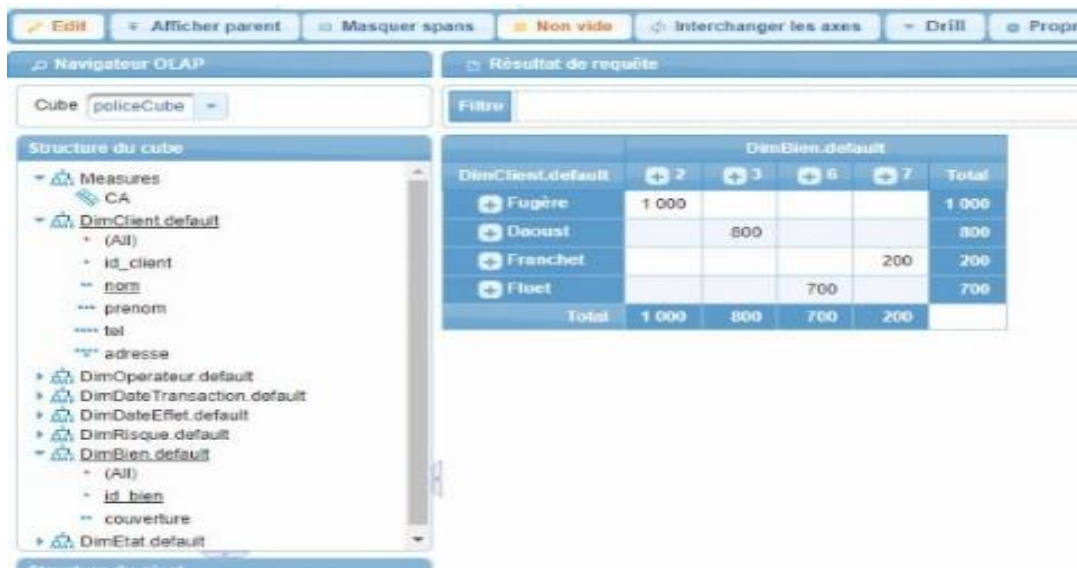


Figure 14 : montant annuel payé par client pour un bien

Chaque ligne correspond à un client et les colonnes contiennent les identifiants des biens.

Exemple : le client Fugère a payé un montant annuel de 1000 pour son bien 2.

On peut aussi analyser le chiffre d'affaires pour chaque bien, ceci est illustré dans la figure 15 :



Figure 15 : chiffre d'affaires pour chaque bien

Le bien dont l'identifiant est 3 a fait un chiffre d'affaires de 800. On peut évaluer ce résultat avec la base de données de départ (figure 16) :

La sélection courante ne contient pas de colonne unique. Les grilles d'édition, les cases à cocher ainsi que les liens Éditer, Copier et Supprimer ne sont pas disponibles.

Affichage des lignes 0 - 5 (total de 6, traité en 0.0000 seconde(s))

SELECT \* FROM "police"

Options

id_dateTransaction	id_dateEffet	id_operateur	id_client	id_etat	id_risque	id_bien	CA
2020-04-19 00:00:00	2020-01-05 00:00:00	1	3	1	2	2	1000
2020-04-19 00:00:00	2020-01-15 00:00:00	1	4	2	3	3	200
2020-04-19 00:00:00	2020-01-15 00:00:00	1	4	2	3	3	300
2020-04-19 00:00:00	2020-01-15 00:00:00	1	4	2	3	3	300
2020-03-10 00:00:00	2020-01-05 00:00:00	2	5	1	7	7	200
2020-01-15 00:00:00	2020-01-05 00:00:00	1	10	1	6	6	700

Opérations sur les résultats de la requête

Imprimer Copier dans le presse-papiers Exporter Afficher le graphique Créer une vue

Figure 16 : table police

Le bien 3 apparaît 3 fois dans la table police, la somme des chiffres d'affaires de ce bien est exactement 800, ce qui correspond exactement au résultat observé dans la figure 15.

Une autre étude peut être réalisée par date. On représente dans la figure 17 les transactions et le chiffre d'affaires réalisé par mois.

Navigateur OLAP

Cube: policeCube

Structure du cube

- Measures
  - DimClient.default
  - DimOperateur.default
  - DimDateTransaction.default
    - id\_date
    - mois
  - DimDateEffet.default
  - DimRisque.default
  - DimBien.default
  - DimEtat.default

Résultat de requête

DimDateTransaction.default	Measures
2020-01-05 00:00:00.0	
2020-01-15 00:00:00.0	700
2020-01-22 00:00:00.0	
2020-03-10 00:00:00.0	200
2020-03-26 00:00:00.0	
2020-04-19 00:00:00.0	1 800
2020-05-13 00:00:00.0	
2020-05-22 00:00:00.0	
2020-12-30 00:00:00.0	

Figure 17 : chiffre d'affaires par date et par bien

L'ensemble de transactions réalisé pendant le mois 4 ont fait un chiffre d'affaires de 1800, chose qu'on peut approuver via la figure 11, en sommant les chiffres d'affaires réalisés pendant le mois en question.

## Conclusion :

Dans cette partie, nous sommes arrivés à créer les cubes ROLAP et les publier sur le serveur Pentaho à l'aide du plugin Pivot4J, ce travail est donc à pour but la création de différents tableaux de bord, facilitant la vue et l'analyse des résultats. C'est le sujet de l'étape suivante.

# Implémentation des tableaux de bord

Les tableaux de bord sont l'outil permettant d'offrir une vue unifiée et d'un format facile à analyser. Ils sont souvent utilisés par les cadres supérieurs pour suivre l'évolution des activités réalisées par leurs entreprises.

Plusieurs outils permettent de réaliser les tableaux de bord, parmi lesquels nous avons utilisé le Power BI.

## Power BI :

Le power BI est un outil développé par Microsoft, pour permettre aux entreprises d'analyser et de visualiser les données en provenance de sources multiples.

Cette solution met l'accent sur le self-service pour permettre à tous les employés de comprendre les données et de les exploiter.

## Tableau de bord

Pour cette agence d'assurance, nous avons implémenter un tableau de bord représentant de différentes statistiques (figure 18).

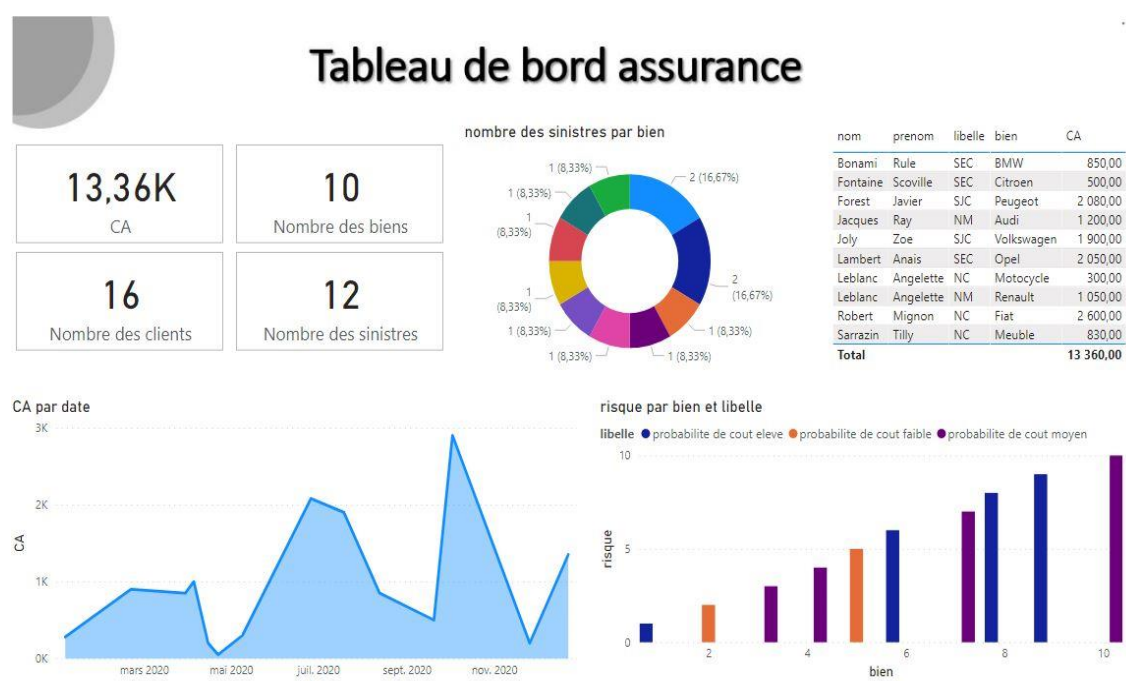


Figure 18 : tableau de bord de l'assurance

Ce tableau de bord permet de visualiser le chiffre d'affaires fait par cette compagnie, nombre de clients existants, nombre de biens et des sinistres.

D'autre part, représente le chiffre d'affaires réalisé pour chaque mois.

La représentation des chiffres d'affaires réalisé par chaque client de la compagnie d'assurance est aussi assurée par ce tableau de bord. Notamment le client Robert a fait 2600 pour son bien Fiat.

Un autre axe d'analyse est celui d'observer les couts des risques fait. On peut citer que le bien 10 avait des risques avec couts élevés.

## **Conclusion générale**

Ce rapport représente toutes les étapes nécessaires pour la conception et la réalisation d'un système décisionnel pour une compagnie d'assurance.

Cela avait pour objectif la réalisation des systèmes facilitant à cette compagnie la prise des bonnes décisions, et l'aide à la bonne analyse des résultats, suivant différents axes de mesures.

La réalisation du travail passant d'une conception d'une base de données sources, vers un entrepôt de données, et puis vers la publication des données sous formes de cubes analysable, a élaboré au final, un résultat sous la forme d'un tableau de bord permettant aux gestionnaires de la compagnie d'assurance, d'avoir une vue claire et nette, afin d'analyser et de développer la compagnie.