

PROJET RÉALISÉ PAR L'ÉQUIPE 04

RAPPORT DE GROUPE EN SCIENCES DES
DONNÉES 2 + BASES DE DONNÉES

MEDOM Michèle Marylyne/ NIYONKURU Berline Cléria/ PALENFO Grace
Sidiqa/ TAHARASTE Rayan



Département MIASHS, UFR 6 Informatique, Mathématique et Statistique
Université Paul Valéry, Montpellier 3

Mai 2023

SOUMIS COMME CONTRIBUTION PARTIELLE
POUR LE COURS SCIENCE DES DONNÉES 2 ET BASES DE DONNÉES

Déclaration de non plagiat

Nous déclarons que ce rapport est le fruit de notre seul travail, à part lorsque cela est indiqué explicitement.

Nous acceptons que la personne évaluant ce rapport puisse, pour les besoins de cette évaluation:

- la reproduire et en fournir une copie à un autre membre de l'université; et/ou,
- en communiquer une copie à un service en ligne de détection de plagiat (qui pourra en retenir une copie pour les besoins d'évaluation future).

Nous certifions que nous avons lu et compris les règles ci-dessus.

En signant cette déclaration, nous acceptons ce qui précède.

Signature: MEDOM Michèle Date: 07/03/2023

Signature: NIYONKURU Berline Date: 07/03/2023

Signature: PALENFO Grace Date: 07/03/2023

Signature: TAHARASTE Rayan Date: 07/03/2023

Table des matières

Chapitre 1	Présentation du thème	1
Chapitre 2	Dénomination des données	2
2.1	Provenance des données	2
2.2	Description des tables	2
2.3	Spécification des sigles	3
Chapitre 3	Modélisation globale des données	5
3.1	Modélisation conceptuelle	5
3.2	Modèle Organisationnel de données	5
Chapitre 4	Prétraitements	7
Chapitre 5	Requêtes réalisées	8
Chapitre 6	Modélisation statistique	13
Chapitre 7	Difficultés et Conclusion	15
7.1	Difficultés	15
7.2	Conclusion	15
Annexes		16
	Graphiques concernant la modélisation en R	16

CHAPITRE 1

Présentation du thème

« **Où rouler en sécurité** » tel est le thème que nous avons choisi d'explorer. Pour aborder ce thème, il nous semble évident d'identifier les circonstances les plus dangereuses sur la route afin d'y être le plus en sécurité possible. Cela nous amène à nous poser alors cette question :

Quelles sont les conditions des routes les plus à risque qui causent des accidents ?

Pour répondre au mieux à la tâche qui nous incombe, nous avons décidé d'assoier une répartition des tâches afin d'optimiser au mieux notre travail et de permettre à chacun des membres de l'équipe d'être à son aise. Cependant cette distribution des tâches ne constitue en aucun cas un étalage des avis sur l'effort par chacun:

MEDOM Michèle Marylyne : Responsable de la manipulations des données,

NIYONKURU Berline Cléria : Responsable de MOD et du MCD,

PALENFO Grace Sidiqa : Responsable de la rédaction des documents,

TAHARASTE Rayan : Responsable du listage des références données et de la partie statistique.

CHAPITRE 2

Dénomination des données

2.1 Provenance des données

Pour répondre à cette problématique, nous allons utiliser des jeux de données ouvertes en licence libre provenant du site [data.gouv.fr]. Les données recueillies ont été préalablement filtrées afin de ne garder que les informations qui nous intéressent. Nous avons deux sources de données.

-Données 1 : Données liées entre elles, portant sur les accidents corporels de la route de 2021 dont la dernière mise à jour date de novembre 2022. On les retrouve dans leur état d'origine en suivant ce lien : {<https://www.data.gouv.fr/fr/datasets/bases-de-donnees-annuelles-des-accidents-corporels-de-la-circulation-routiere-annees-de-2005-a-2020/>}

-Données 2 : Données décrivant les réseaux cyclables en novembre 2022, disponible par ce lien ci : {<https://www.data.gouv.fr/fr/datasets/reseau-des-itinéraires-cyclables/>}.

De façon générale, le travail sur les données portera uniquement que le jeux de données 1. Nous avons enlevé les colonnes qui ne nous semblaient pas pertinentes, les variables qui n'affectent pas les chances d'être impliquées dans un accident sur la route lorsque l'on est dans un véhicule. Cela a permis de dégager trois (3) tables pour la suite du projet à savoir **Accidents**(6 colonnes,25284 lignes), **Lieux**(8 colonnes,25284 lignes) et **Véhicules**(4 colonnes,43392 lignes).

2.2 Description des tables

-La table **Accidents** : Dans cette table, nous avons gardé les colonnes décrivant le numéro d'identifiant de l'accident (Num_Acc) unique de la table qui permet de lier cette table aux autres tables, le département (dep), la commune (com), le type d'intersection (int), les conditions atmosphériques (atm), le type de collision (col). Le Num_Acc qui est un Integer unique correspond à la clé primaire. On a aussi ajouté id_lieu(Integer) comme une clé étrangère.

-La table **LIEUX** : Dans cette table, un identifiant unique qu'on a nommée id_lieu(Integer) a été créé pour servir de lien entre cette table et les autres. Elle constitue la clé primaire. Nous avons gardé les colonnes décrivant le numéro d'identifiant de l'accident(Num_Acc), la catégorie de route(catr), le régime de circulation(circ), le nombre de voies(nbv), l'existence d'une voie réservée(vosp), l'état de la surface de la route(surf),les infrastructures comme un pont ou une voie ferrée(infra)et la vitesse maximale autorisée au moment de l'accident(vma).

-La table **VEHICULES** : Dans cette table, sont retenues les colonnes décrivant le numéro d'identifiant de l'accident (Num_Acc), l'identifiant unique du véhicule (id_vehicule), l'identifiant du lieu de l'accident (id_lieu) et la catégorie du véhicule (catv). La clé primaire est l'id_vehicule.

Nom colonne	Type	Caractéristique
Accidents	table	6 colonnes, 25284 lignes
Lieux	table	8 colonnes, 25284 lignes
Vehicules	table	4 colonnes, 43392 lignes

2.3 Spécification des sigles

Vous trouverez ci dessous la définition des abréviations des noms et des sigles utilisés tout le long de ce projet.

ACCIDENTS :

Num_Acc : numéro d'identifiant de l'accident, entier, unique, pas de valeur manquante, clés de la table.

id_lieu : numéro d'identifiant unique du lieu d'accident véhicule.

Dep : le département où a eu lieu l'accident, entier, non unique, pas de valeur manquante.

Com : la commune où a eu lieu l'accident, entier, non unique, pas de valeur manquante.

Agg : si l'accident a eu lieu en ou hors agglomération, entier, non unique, pas de valeur manquante.

Atm : les conditions atmosphériques lors l'accident, entier, non unique, pas de valeur manquante.

Col : le type de collisions , entier, non unique, pas de valeur.

LIEUX :

Num_Acc : numéro d'identifiant de l'accident, entier, unique, pas de valeur manquante, clés de la table.

id_lieu : numéro d'identifiant unique du lieu d'accident véhicule.

Catr : catégorie de la route, entier, non unique, pas de valeur.

Nbv : nombre de voies de circulation, entier, non unique, pas de valeur manquante.

Vosp : indique existence d'une voie réservée et son type si elle existe, entier, non unique, pas de valeur manquante.

Surf : état de la surface de la route, entier, non unique, pas de valeur.

Infra : présence et type d'infrastructures tels un tunnel ou une zone de péage, entier, non unique, pas de valeur manquante.

Vma : vitesse maximale autorisée sur le lieu de l'accident au moment de ce dernier, entier, non unique, pas de valeur manquante.

VEHICULES :

Num_Acc : numéro d'identifiant de l'accident, entier, unique, pas de valeur manquante.

id_vehicule : numéro d'identifiant unique du véhicule, entier, unique, pas de valeur manquante.

id_lieu : numéro d'identifiant unique du lieu d'accident véhicule.

Catv : catégorie du véhicule, entier, unique, pas de valeur.

CHAPITRE 3

Modélisation globale des données

3.1 Modélisation conceptuelle

Après avoir trié, filtré et fait le grand « ménage » dans les données, nous avons pu dresser un modèle conceptuel adapté à nos différentes tables grâce à l'outil en ligne [<https://mocodo.net/>] . Figure ci dessous:

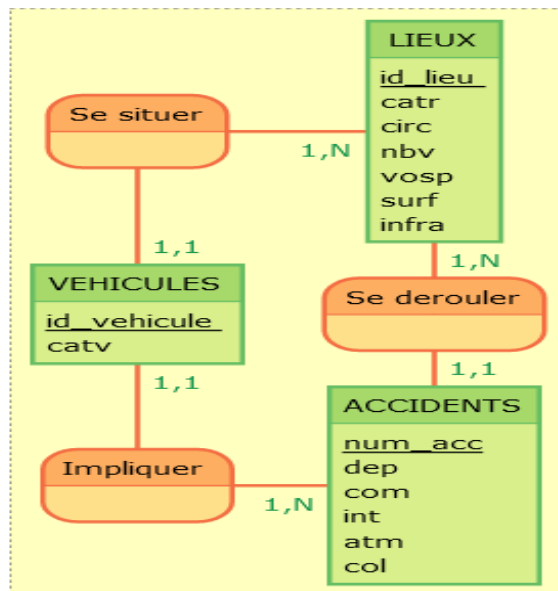


Figure 3.1: MCD

3.2 Modèle Organisationnel de données

La version manuscrite du MOD ci dessous :

LIEUX (id_lieu, num_acc, catr, circ, nbv, vosp, surf, infra)
VEHICULES (id_vehicule, num_acc, id_lieu, catv)
ACCIDENTS (num_acc, id_lieu, dep, com, int, atm, col)

À partir du Modèle Conceptuel des Données, nous avons créé le Modèle Organisationnel des données, grâce à l'outil Concepteur de MAMP - PhpMyAdmin:

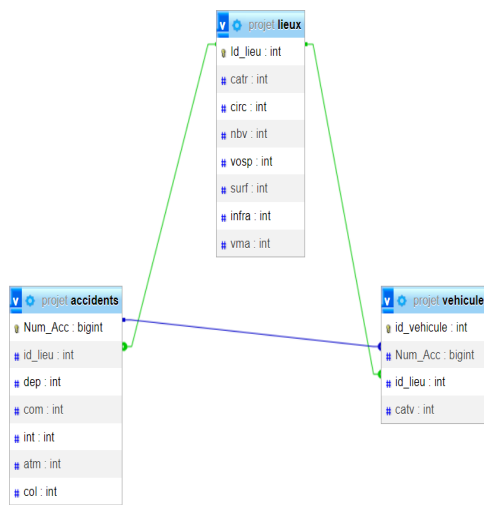


Figure 3.2: mod

CHAPITRE 4

Prétraitements

Afin de manipuler aisément les données dans SQL, nous les avons fait liés via les clés primaires et étrangères cités auparavant. Ensuite, pour faciliter l'import de nos données sur la plateforme en ligne [<https://filess.io/>], nous avons dû modifier les tables dans les différents classeurs Excel.

-Pour la table **Accidents** : nous ne gardons pas les colonnes décrivant le jour(jour), le mois(mois), l'heure à la minute près(hrmn), la lumière au moment de l'accident(lum),l'année(an) car toutes les données datent de la même année et l'adresse postale(adr) car la valeur n'est pas renseignée pour tous les accidents donc moins pertinents.

-Pour la table **Lieux** : Ont été exclus les colonnes décrivant l'inclinaison de la route(prof), la largeur de la chaussée affectée à la circulation hors bandes d'arrêt d'urgences et places de stationnement ainsi que terre-plein-central (larROUT), le numéro de la route(voie), l'indice numérique de route(V1), la lettre indice alphanumérique de la route(V2), le numéro du PR de rattachement(pr), la distance en mètres au PR(pr1), le tracé en plan(plan) ainsi que la largeur du terre-plein central(TPC) car cela ne nous semblaient pas être des variables primordiales influençant énormément sur les chances d'être impliqués dans un accident.

-Pour la table **Vehicules** : Nous avons exclus les colonnes décrivant le point de choc initial car ces dernières n'avaient absolument aucun impact sur les chances d'être impliqués dans un accident(choc), le type de motorisation du véhicule (motor) et le nombre de personne dans le véhicule (occutc),le sens de circulation (senc), l'obstacle fixe heurté si c'est le cas (obs), l'obstacle mobile heurté si c'est le cas (obsM) ainsi que la manoeuvre effectué avant l'accident(many).

CHAPITRE 5

Requêtes réalisées

```
# install.packages("RMySQL")
# install.packages("DBI")
library(DBI)

## Warning: le package 'DBI' a été compilé avec la version R 4.2.3

con <- DBI::dbConnect(
  drv = RMySQL::MySQL(),
  host = "nkv.h.files.io",
  port = 3307,
  username = "Projet_produceegg",
  password = "5de5b1c29e6b2612ba58fcede731030b73ca5cf8",
  dbname = "Projet_produceegg"
)

show tables;
```

Table 5.1: 3 records

Tables_in_Projet_produceegg
accidents
lieux
vehicules

Requête 1: *Lister le nombre d'accidents par département.*

Le département 75, celui de *Paris* cumule le plus d'accidents. Nous avons décidé d'ajouter la ligne "having nombre ≥ 1000 " par soucis d'affichage et ainsi on obtient *le Graphe1* qui est un peu trop surchargé mais on y voit clairement les départements les plus impliqués dans des accidents.

```
SELECT dep, COUNT(DISTINCT accidents.Num_Acc) as nombre
FROM accidents
GROUP BY dep
HAVING nombre  $\geq 1000$ 
ORDER BY nombre DESC
```

Table 5.2: 6 records

dep	nombre
75	2133
93	1240
13	1175
94	1095

dep	nombre
92	1081
69	1062

Requête 2: *Quelles sont les catégories de véhicule les impliquées dans les accidents?*

La catégorie de véhicules la plus impliquée est la n°7 qui correspond au **Véhicule Léger**.

```
SELECT vehicules.catv, COUNT(DISTINCT accidents.Num_Acc) as accident
FROM vehicules,accidents,lieux
WHERE accidents.id_lieu=lieux.Id_lieu
AND vehicules.Num_Acc=accidents.Num_Acc
GROUP BY vehicules.catv
ORDER BY accident DESC;
```

Table 5.3: Displaying records 1 - 10

catv	accident
7	19191
33	3098
10	2822
1	2163
2	1592
30	1393
32	967
50	752
31	699
15	415

Requête 3: *Quels sont les types de collisions les plus fréquents lors des accidents ?*

Les collisions impliquants *deux voitures par le coté* sont les plus fréquentes lors des accidents.

```
SELECT col, COUNT(*)
FROM accidents
GROUP BY col
ORDER BY COUNT(*) DESC
```

Table 5.4: 8 records

col	COUNT(*)
3	7779

col	COUNT(*)
6	7219
2	3388
1	2646
7	2306
4	932
5	725
-1	5

Requête 4: *Quels sont les infrastructures/aménagements les plus touchés par les accidents?*

Dans plus de la moitié des accidents aucune infrastructure n'est touché.

```
SELECT lieux.infra, COUNT(DISTINCT accidents.Num_Acc) as count
FROM accidents, lieux
WHERE accidents.id_lieu=lieux.id_lieu
GROUP by lieux.infra
ORDER BY count DESC
```

Table 5.5: Displaying records 1 - 10

infra	count
0	21025
5	1253
9	904
2	373
3	368
1	310
6	249
-1	247
8	161
4	98

Requête 5: *Quels états de la surface de la route sont les plus dangereux?*

Quasiment tous les accidents ont lieu quand sur des surfaces de *routes normales* suivis de près par les accidents sur des surfaces *mouillées*.

```
SELECT lieux.surf, COUNT(DISTINCT accidents.Num_Acc)
FROM accidents, lieux
WHERE accidents.id_lieu=lieux.Id_lieu
GROUP BY lieux.surf
ORDER BY `COUNT(DISTINCT accidents.Num_Acc)` DESC
```

Table 5.6: Displaying records 1 - 10

surf	COUNT(DISTINCT accidents.Num_Acc)
1	19758
2	4826
9	155
7	117
3	39
6	33
8	30
5	26
-1	10
4	6

Requête 6: *Sous quelles vitesses maximales autorisées (vma) sur les routes sont répertoriées les accidents les plus à risques.*

Les accidents ont plutôt tendance à avoir lieu sur des voies à 50 Km/h.

```
SELECT lieux.vma, COUNT(DISTINCT accidents.Num_Acc)
FROM accidents, lieux
WHERE accidents.id_lieu=lieux.Id_lieu
GROUP BY lieux.vma
ORDER BY `COUNT(DISTINCT accidents.Num_Acc)` DESC
```

Table 5.7: Displaying records 1 - 10

vma	COUNT(DISTINCT accidents.Num_Acc)
50	12330
30	3454
80	3362
90	2021
70	1810
110	907
130	470
-1	414
20	61
10	47

Requête 7: *Lister le nombre de voies (nbv) selon les accidents les plus fréquents*

Ces accidents sont répertoriés le plus souvent sur les routes à 2 voies.

```
SELECT lieux.nbv, COUNT(DISTINCT accidents.Num_Acc)
FROM accidents, lieux
WHERE accidents.id_lieu=lieux.Id_lieu
```

```
GROUP BY lieux.nbv
ORDER BY `COUNT(DISTINCT accidents.Num_Acc)` DESC;
```

Table 5.8: Displaying records 1 - 10

nbv	COUNT(DISTINCT accidents.Num_Acc)
2	15162
4	3024
1	2667
3	2010
6	658
0	547
5	378
-1	221
8	208
7	54

Requête 8: *Quelles sont les conditions atmosphériques (atm) ayant causées plus d'accidents?*

Quasiment tous les accidents ont lieu quand les conditions atmosphériques sont 1: normale et 2: mouillées.

```
SELECT atm,COUNT(DISTINCT accidents.Num_Acc) as nombre
FROM accidents
GROUP BY atm
ORDER BY nombre DESC
```

Table 5.9: Displaying records 1 - 10

atm	nombre
1	19588
2	2782
8	1074
3	521
7	447
5	337
9	143
4	58
6	43
-1	7

CHAPITRE 6

Modélisation statistique

Pour faciliter la manipulation des requêtes avec Rstudio, nous avons dû ajuster l'écriture des requêtes qui restent dans le fond les mêmes avec des résultats similaires.

Requête 1 #Graphe1 : *Porte sur la fréquence des accidents, requête 1 en SQL.*

On peut constater que le département ayant le plus d'accident est le département de Paris et ceux d'après sont les départements autour de Paris (93,94 et 92) à l'exception du 13 qui est les Bouches-du-Rhône où se trouve Marseille. Ces résultats ne sont pas si surprenant car ce sont sûrement les départements où il y a le plus de circulation.

Requête 2 #Graphe2 : *Porte sur les catégories de véhicules impliquées dans les accidents, requête 2 en SQL.*

La requête nous a permis de voir que les catégories de véhicules les plus impliquées dans les accidents sont les véhicules de type 7 et 33 selon les résultats de notre requête sql. Les véhicules de type 7 sont des véhicules légers, et donc regroupe tous les véhicules de moins de 3,5 tonnes (les voitures) et les véhicules de type 33 sont les motocyclettes.

Les motocyclettes et les voitures légers selon le camembert obtenu à partir de la requête faite sur R sont les véhicules qui enregistrent le plus grand nombre d'accidents; ce qui est logique car ce sont les véhicules les plus utilisés sur la route.

Requête 3 #Graphe3 : *Porte sur les collisions, requête 3 en sql.*

Nous avons fait une requête afin d'analyser les types de collisions les plus fréquents dans les départements les plus à risques. Cette requête sql nous a donné des résultats selon les 8 types de collisions. Mais nous nous intéressons aux deux types de collisions qui regroupent le plus grand nombre d'accidents. Nous voyons que les collisions de types 3 et 6 sont les plus fréquentes. Le type 3 désigne la collision de deux véhicules par le côté et le type 6 désigne toutes les collisions simples qui n'incluent qu'un seul véhicule.

Nous avons comparé les résultats des types collisions les plus fréquents dans tous les départements aux résultats des types de collisions les plus fréquents dans les régions les plus à risques et avons bien remarqué sur les camemberts qu'ils sont similaires. La seule différence notable est que les collisions de type 3 sont un peu plus fortes dans les départements les plus à risques.

Requête 4 #Graphe4 : *Porte sur les surfaces, requête 5 en sql.*

En outre, nous avons fait cette requête afin de connaître les états de la route les plus dangereuses. Selon les résultats de notre requête sql ce sont les états de surface

de la route de type 1 et 2 qui causent le plus d'accidents. Celui de type 1 représente lorsque la route a un état normal, c'est à dire que le sol n'est ni enneigé ou mouillé, il y a rien sur la route. L'état de surface de type 2, c'est lorsque le sol est mouillé. Nous remarquons qu'en situation de surface normale, il y'a beaucoup plus d'accidents que la surface des routes est mouillée, lorsqu'il pleut. Cependant, nous aurons tendance à penser que c'est l'inverse qui se produit mais ce n'est pas le cas en réalité. C'est certainement dû au fait qu'il y a plus de circulation lorsque le sol n'est pas mouillée que lorsqu'il l'est.

Requête 5 #Graphe5 : *Porte sur la vitesse maximale, requête 6 en sql.*

On constate que dans la plupart des accidents la vitesse maximale autorisée est de 50 km/h. On peut voir que c'est peu près la moitié des accidents qui sont concernés et qui se produisent donc très certainement en agglomération. Et juste en dessous la seconde vitesse maximale autorisée impliquée dans le plus d'accident est 30 km /h. Ainsi, on peut en déduire que la majorité des accidents se produisent en agglomération ce qui n'est pas très étonnant car il est plus compliqué de rouler en agglomération que sur la route hors agglomération. On peut voir qu'un bon quart des accidents on l'air de se dérouler sur des départementales et le reste probablement sur l'autoroute.

Requête 6 #Graphe6 : *Porte sur les conditions atmosphériques les plus fréquentes, requête 8 en sql.*

On remarque assez facilement que la grande majorité des accidents ont lieu pour la condition atmosphérique 1 qui correspond à des conditions atmosphériques normales. On peut donc en conclure que ces dernières ont peu d'impact sur les accidents. On remarque tout de même que les conditions atmosphériques 2 et 8, qui correspondent respectivement à de la pluie légère et à un temps couvert se démarquent un peu mais cela peut s'expliquer par le fait qu'il va certainement avoir plus de gens qui roulent lorsqu'il pleut ou qu'il fait couvert que lorsqu'il neige ou qu'il y a une tempête.

CHAPITRE 7

Difficultés et Conclusion

7.1 Difficultés

Au cours de la réalisation de ce projet, nous avons rencontré certes peu de difficulté mais les quelques difficultés que nous avons pu rencontrer nous ont donné du fil à retordre.

La principale difficulté, qui est aussi la seule notable, que nous avons rencontrés a été la mise en ligne de notre base de données, pour faire ceci il fallait que notre base de données ait une taille inférieure à 10 Mb et à chaque fois que nous importions une base de données en ligne elle était acceptée pour ensuite être désactivé quelques jours plus tard. La raison était que la vraie taille de la base de données mise en ligne mettait plusieurs jours à être calculé et ne correspondais pas exactement à la taille qu'on avait. Nous avons finalement réussi à la mettre en ligne après l'avoir filtré plusieurs fois.

Nous avons aussi rencontré une légère difficulté sur R pour faire les graphiques car juste recopier les requêtes SQL ne fonctionner pas, il a fallu les modifier pour pouvoir faire les graphiques que l'on voulait.

7.2 Conclusion

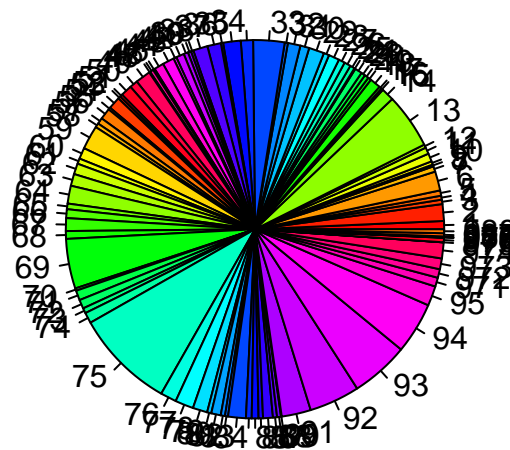
On remarque que les conditions les plus à risques ne sont pas vraiment celles impliqués le plus dans les accidents et qu'au contraire ce sont pour les conditions les moins à risque qu'il y a le plus d'accident. Cela s'explique par le fait qu'il doit simplement y avoir plus de gens qui roule lorsque les conditions sont moins à risque ce qui fait qu'en fin de compte ce n'est pas forcément moins dangereux car comme il y a plus de monde qui en circulation, il y a plus de chance d'être impliqué dans un accident.

Annexes

Graphiques concernant la modélisation en R

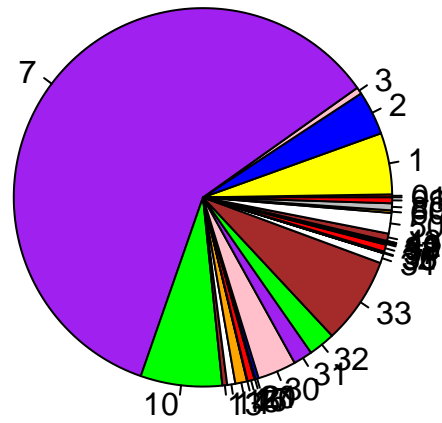
Graphe 1

```
Z<-dbGetQuery(con,"SELECT accidents.dep  
FROM accidents,lieux  
WHERE accidents.id_lieu=lieux.Id_lieu;")  
pie(table(Z), col = rainbow(50))
```



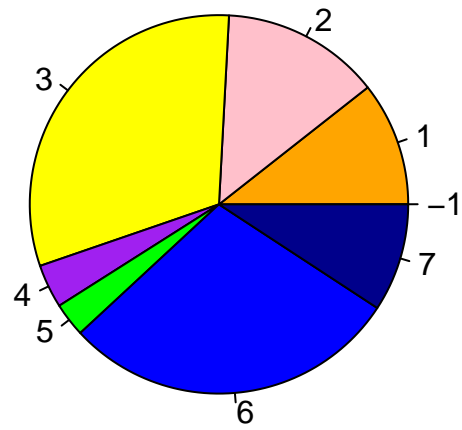
Graphe 2

```
Z<-dbGetQuery(con,"SELECT vehicules.catv  
FROM vehicules,accidents,lieux  
WHERE accidents.id_lieu=lieux.Id_lieu AND vehicules.Num_Acc=accidents.Num_Acc  
")  
pie(table(Z), col = c("gray", "red", "yellow", "blue", "pink", "purple",  
"green", "brown", "white", "orange"))
```

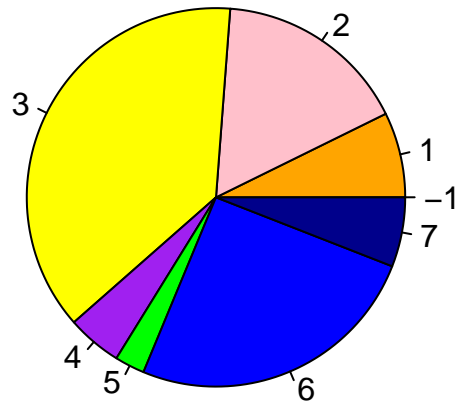


Grape 3

```
Z<-dbGetQuery(con,"SELECT col
FROM accidents")
pie(table(Z), col = c("red","orange","pink", "yellow","purple",
"green","blue", "darkblue"))
```

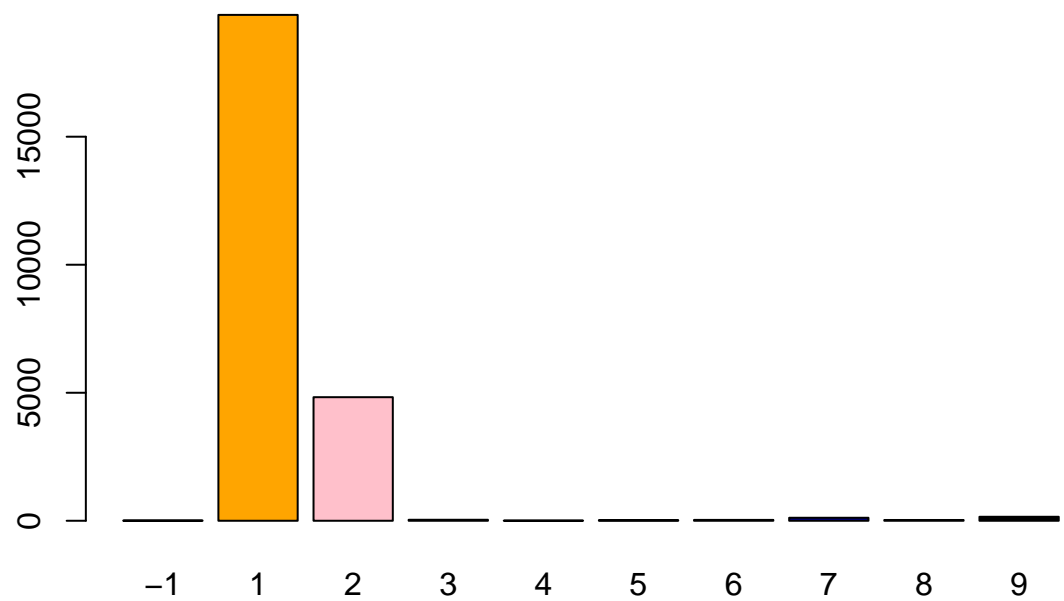


```
#graphe de la répartition du type de collisions dans les départements
#les plus à risques
Z<-dbGetQuery(con,"SELECT col
FROM accidents
where dep=13 OR dep=94 OR dep=93 OR dep=92 OR dep=75 OR dep=69
")
pie(table(Z), col = c("red","orange","pink", "yellow","purple",
"green","blue", "darkblue"))
```



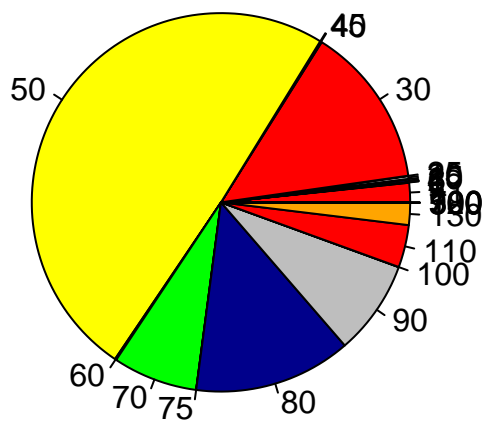
Graphe 4:

```
Z<-dbGetQuery(con,"SELECT lieux.surf FROM accidents,
                    lieux WHERE accidents.id_lieu=lieux.Id_lieu")
barplot(table(Z), col = c("red","orange","pink", "yellow",
                          "purple", "green","blue", "darkblue","gray","black"))
```



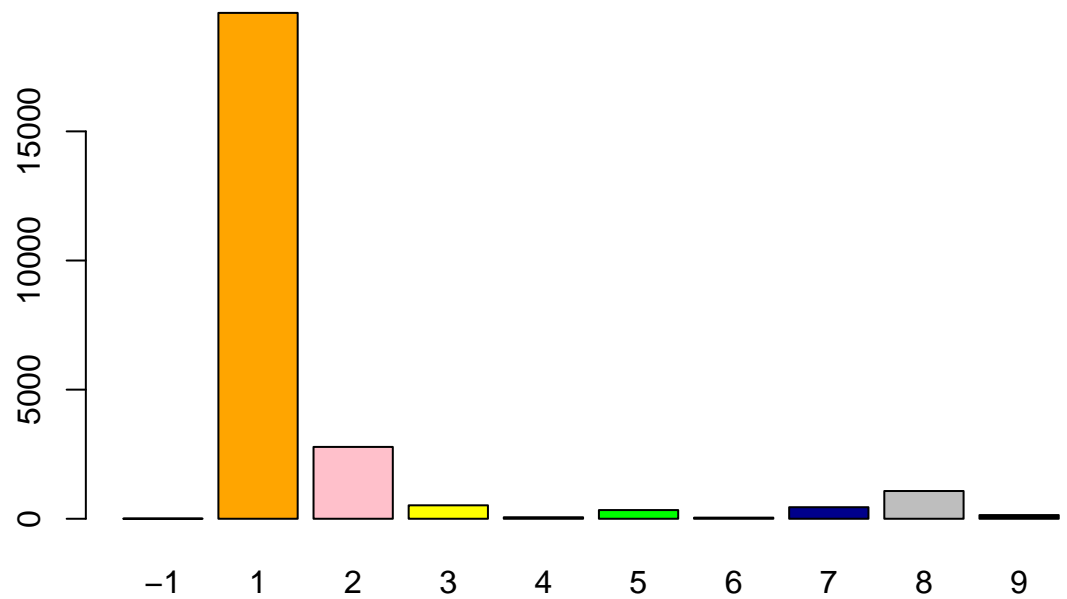
Graphe 5

```
Z<-dbGetQuery(con,"SELECT lieux.vma FROM accidents,
lieux WHERE accidents.id_lieu=lieux.Id_lieu
")
pie(table(Z), col = c("red","orange","pink", "yellow","purple",
"green","blue", "darkblue","gray","black"))
```



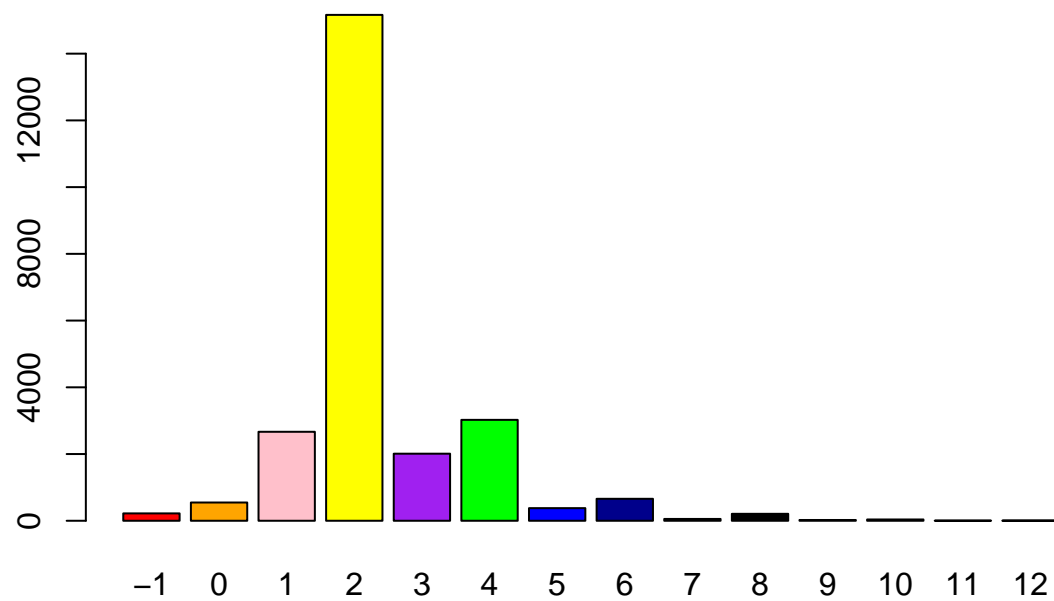
Graphe 6

```
Z<-dbGetQuery(con,"SELECT atm
FROM accidents")
barplot(table(Z), col = c("blue","orange","pink", "yellow","purple", "green",
"blue", "darkblue","gray","black"))
```

Graphe 7

```
Z<-dbGetQuery(con,"SELECT lieux.nbv FROM accidents,
lieux WHERE accidents.id_lieu=lieux.Id_lieu ;
")
barplot(table(Z), col = c("red","orange","pink", "yellow","purple", "green"
, "blue", "darkblue","gray","black"))
```



```
dbDisconnect(con)
```

```
## [1] TRUE
```