

Projet A3

Big-Data

Nicolas Capelle, Thomas Catros, Hermine Paul

Encadrants: Mr. Vignaud, Mr.Saumard

26 - 05 - 2023

Présentation

- Gestion de projet
 - Gantt
- Étapes du projet
 - Nettoyage des données
 - Les graphiques
 - Equations de droite
 - Les cartes
 - Relations entre variables qualitatives

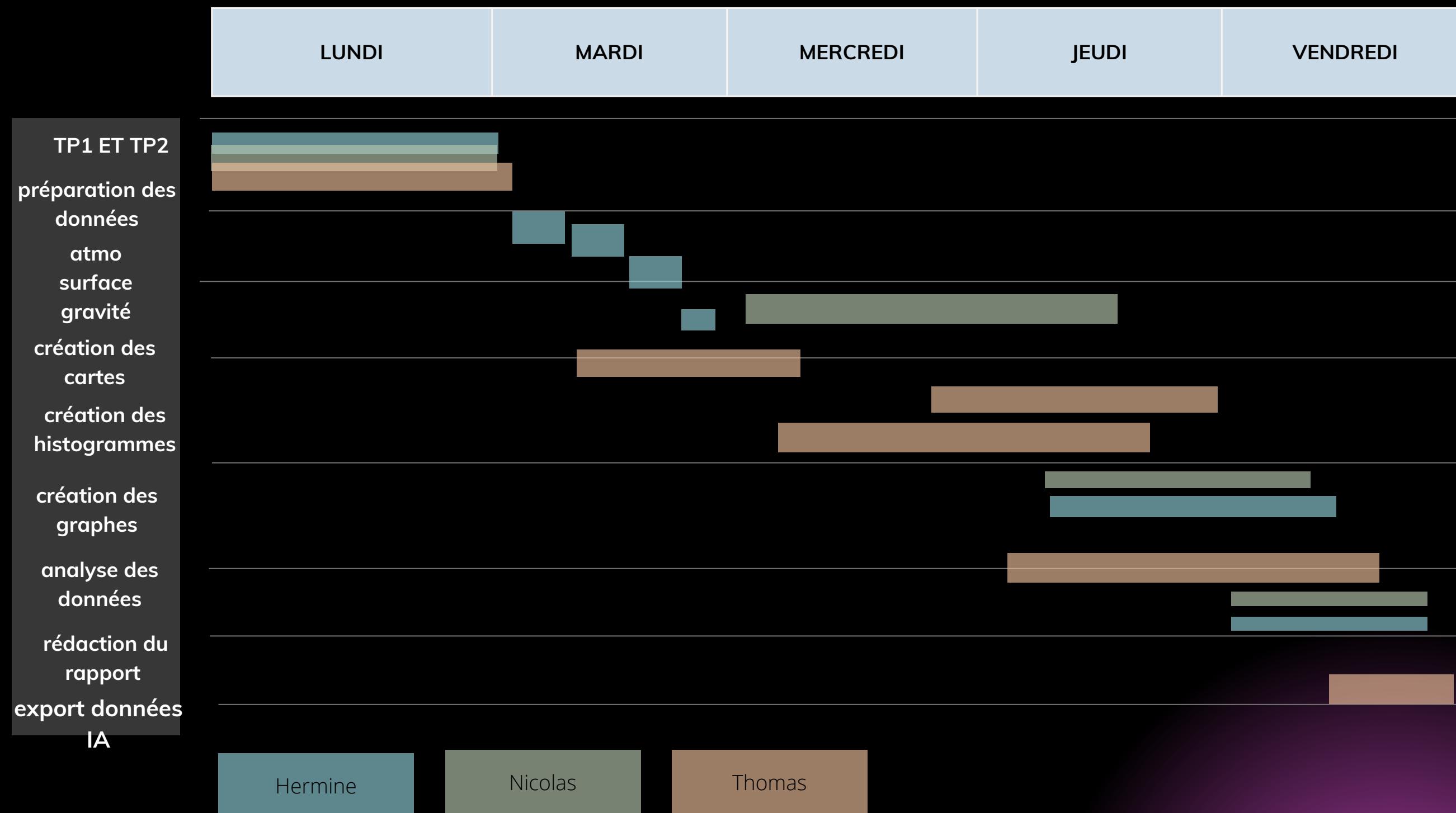


Gestion de projet

Gantt

ISEN
GANTT

OUEST



Nettoyage des données

descr_cat_veh	descr_aggro	descr_athmo	descr_lum	descr_etat_surf	description_intersection
VL seul	En agglomération	Normale	Plein jour	Normale	Intersection en X
VL seul	En agglomération	Normale	Plein jour	Normale	Intersection en X
PL > 3,5T + remorque	En agglomération	Normale	Plein jour	Normale	Hors intersection
Scooter < 50 cm3	Hors agglomération	Normale	Nuit sans éclairage public	Normale	Hors intersection
VL seul	En agglomération	Normale	Nuit avec éclairage public allumé	Normale	Intersection en T
VL seul	Hors agglomération	Normale	Plein jour	Normale	Hors intersection



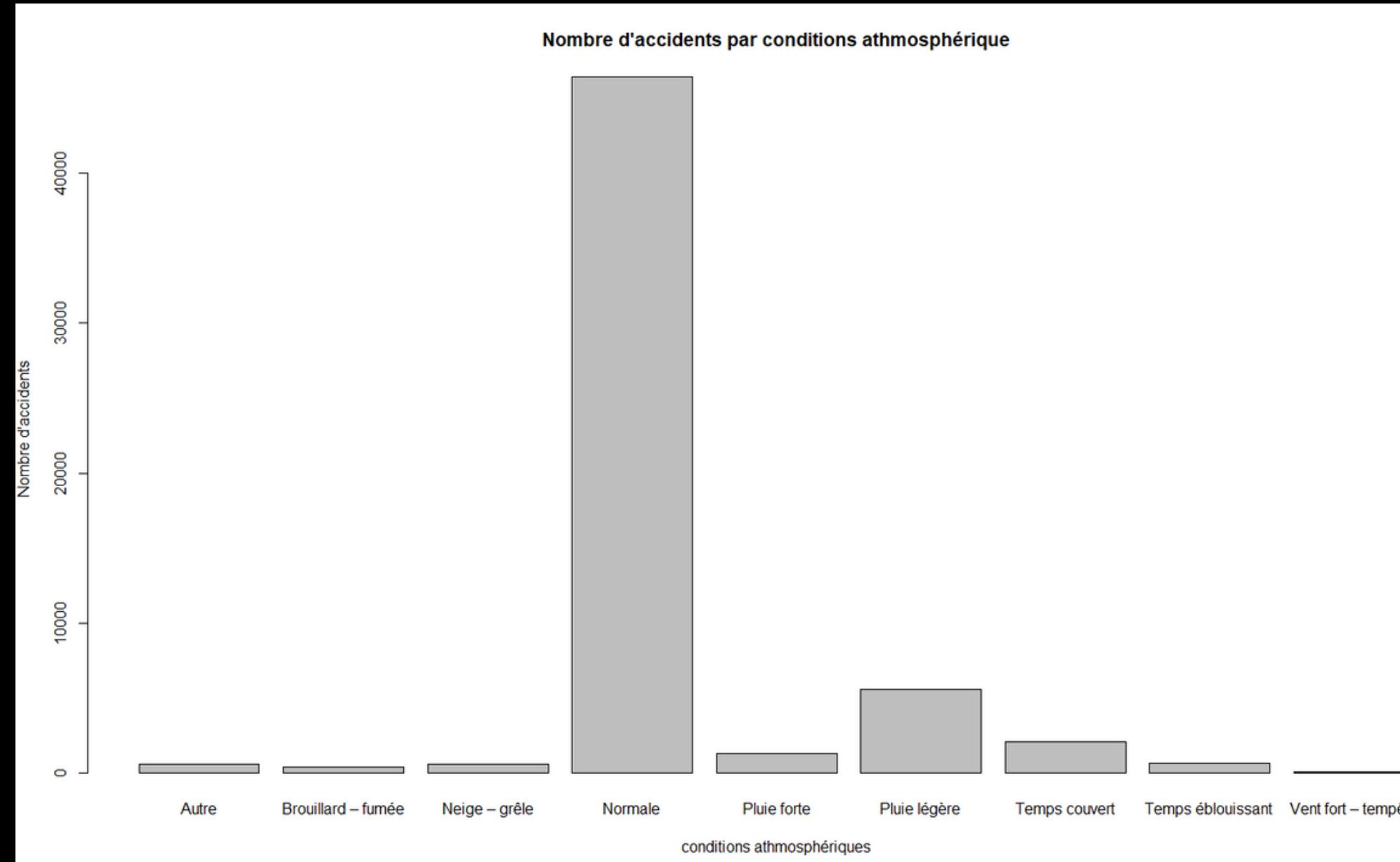
longitude	descr_cat_veh	descr_aggro	descr_athmo	descr_lum	descr_etat_surf	description_intersection	an_nais	age	place	descr_dispo_secu	date
-1.233330	1	0	1	1	1		1	1969	40	1	1
5.633330	2	0	2	2	2		1	1959	50	1	1
5.633330	3	0	2	2	2		1	1960	49	1	1
2.983330	3	0	3	3	3		1	1957	52	1	1
0.066667	3	0	4	2	4		2	1951	58	1	1
-4.050000	2	0	4	2	1		1	1969	40	1	1
3.397230	3	0	4	2	4		3	1978	31	1	1
-3.783330	3	0	5	2	1		1	1949	60	1	1
-0.350000	3	0	2	3	2		1	1967	42	1	1
-0.633333	3	0	5	3	1		1	1971	38	1	1

Graphiques

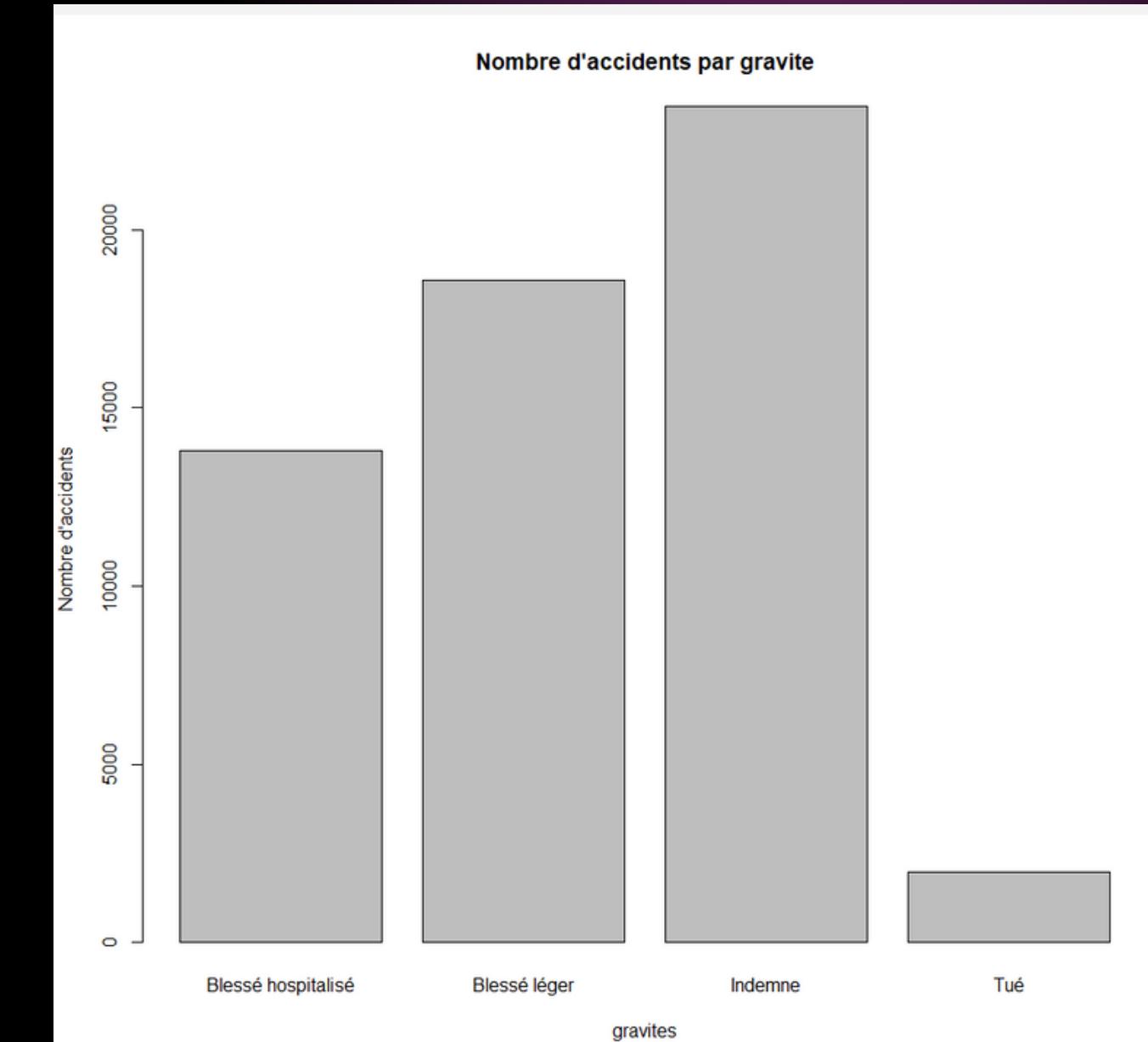
- o Nombre d'accidents en fonction des conditions **atmosphériques**
- o Nombre d'accidents en fonction de la description de la **surface**
- o Nombre d'accidents selon la **gravité**
- o Nombre d'accidents par **tranches d'heure**
- o Nombre d'accidents par **ville**



Graphiques

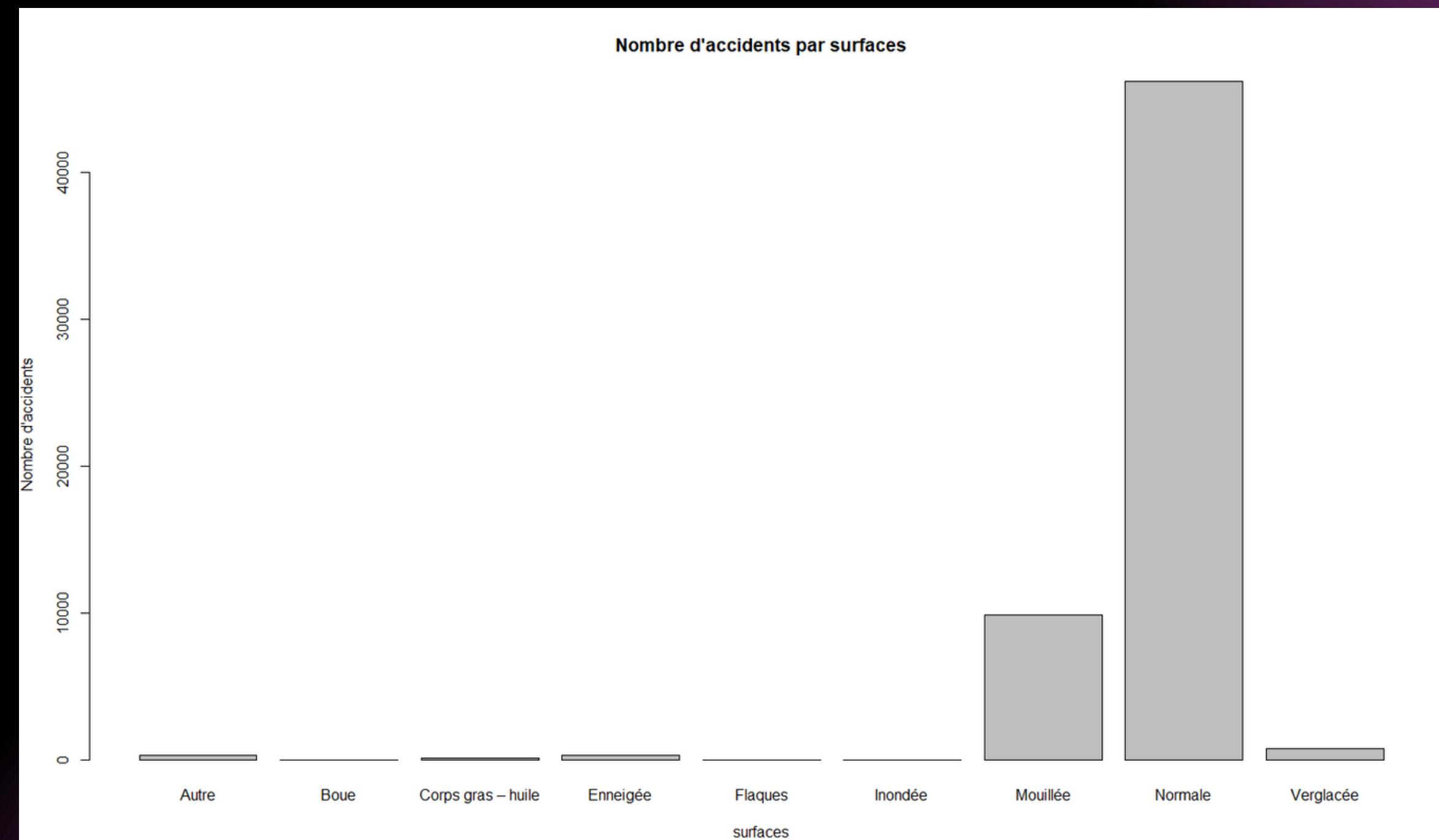


Nombre d'accidents en
fonction des conditions
atmosphériques

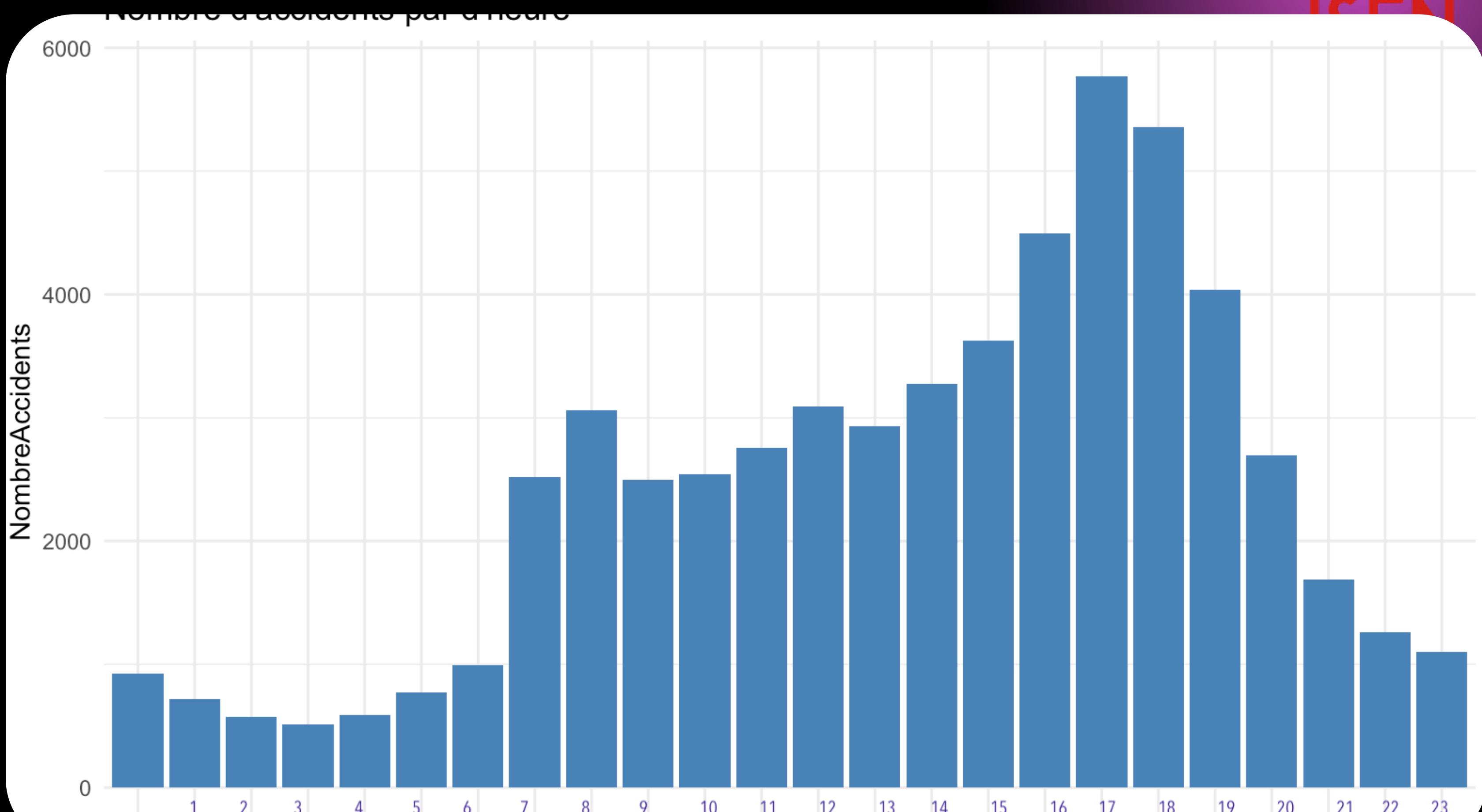


Nombre d'accidents
selon la **gravité**

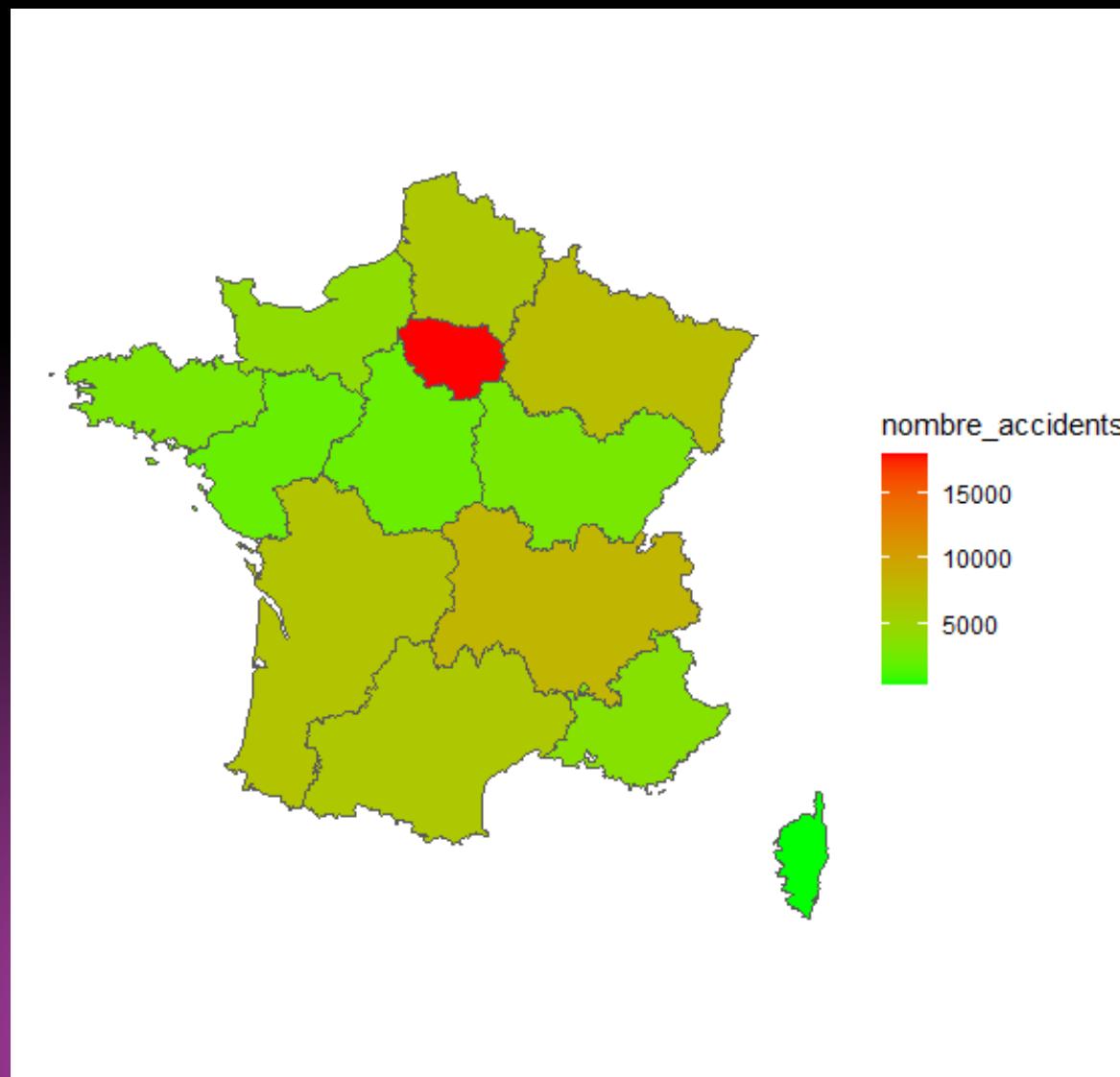
Graphiques



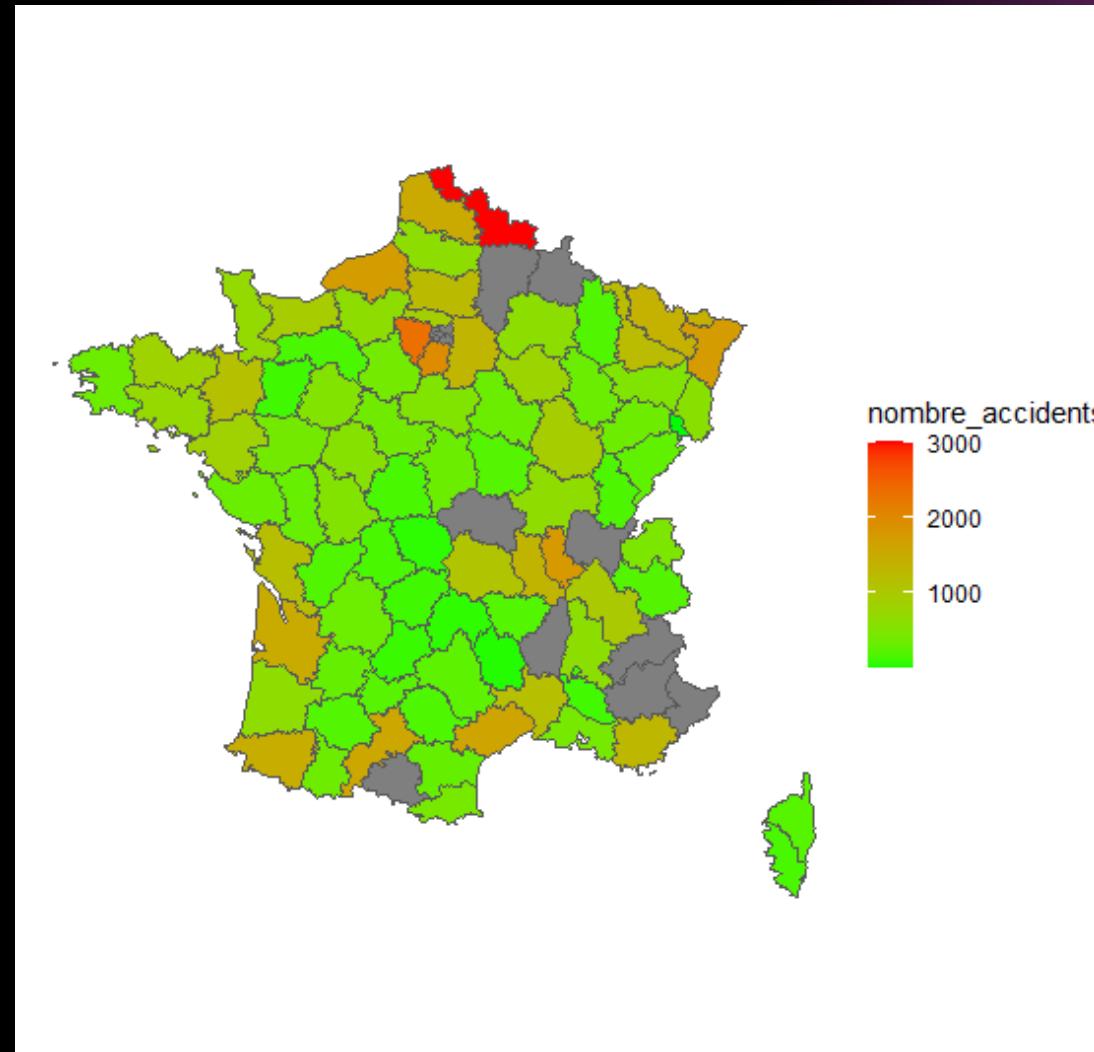
Nombre d'accidents en fonction
de la description de la **surface**



Cartes régions et départements



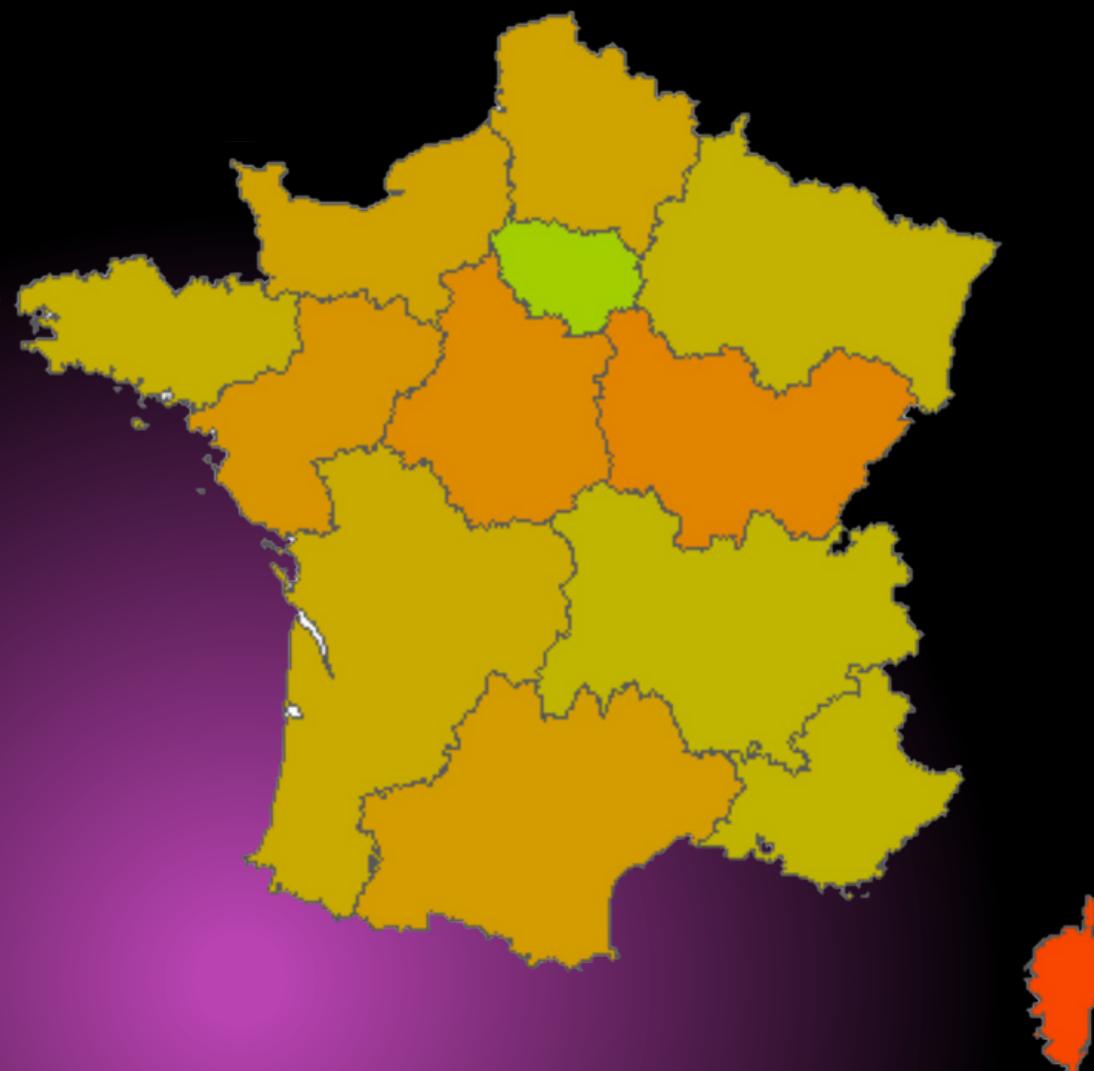
Nombre
d'accidents par
région



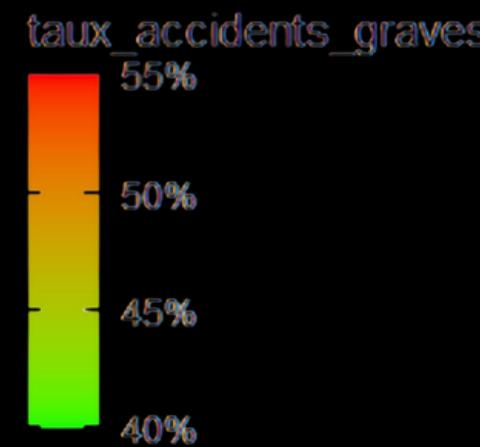
Nombre
d'accidents par
département

Cartes

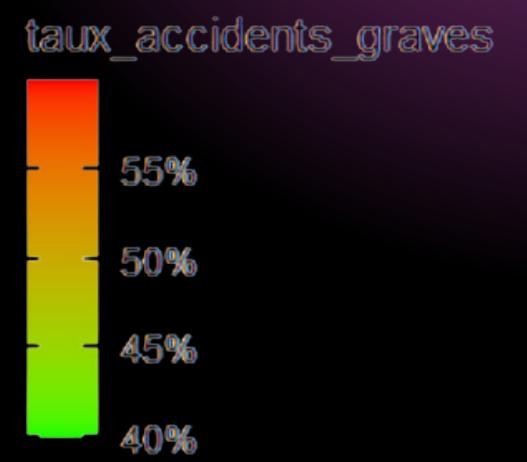
régions et
départements



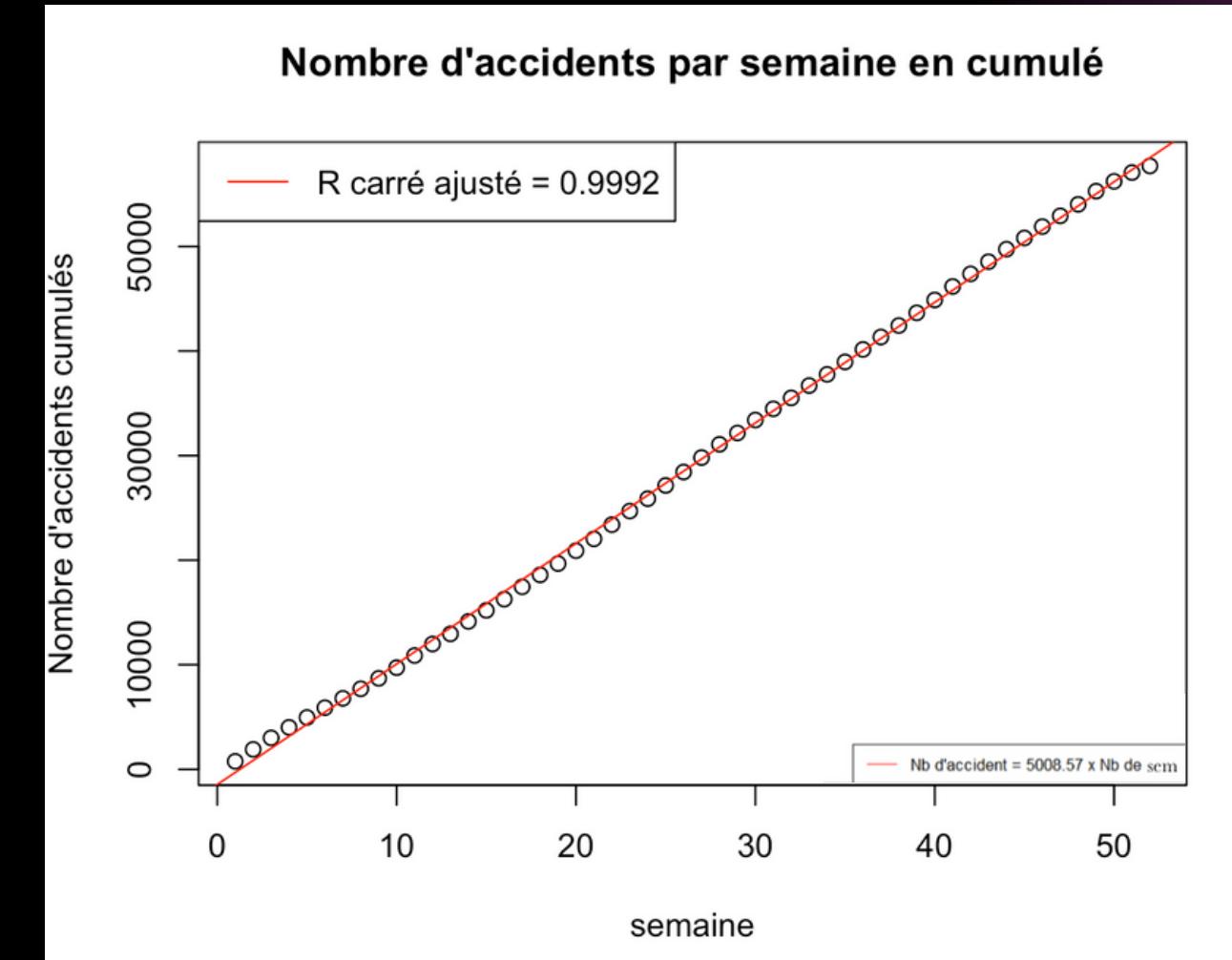
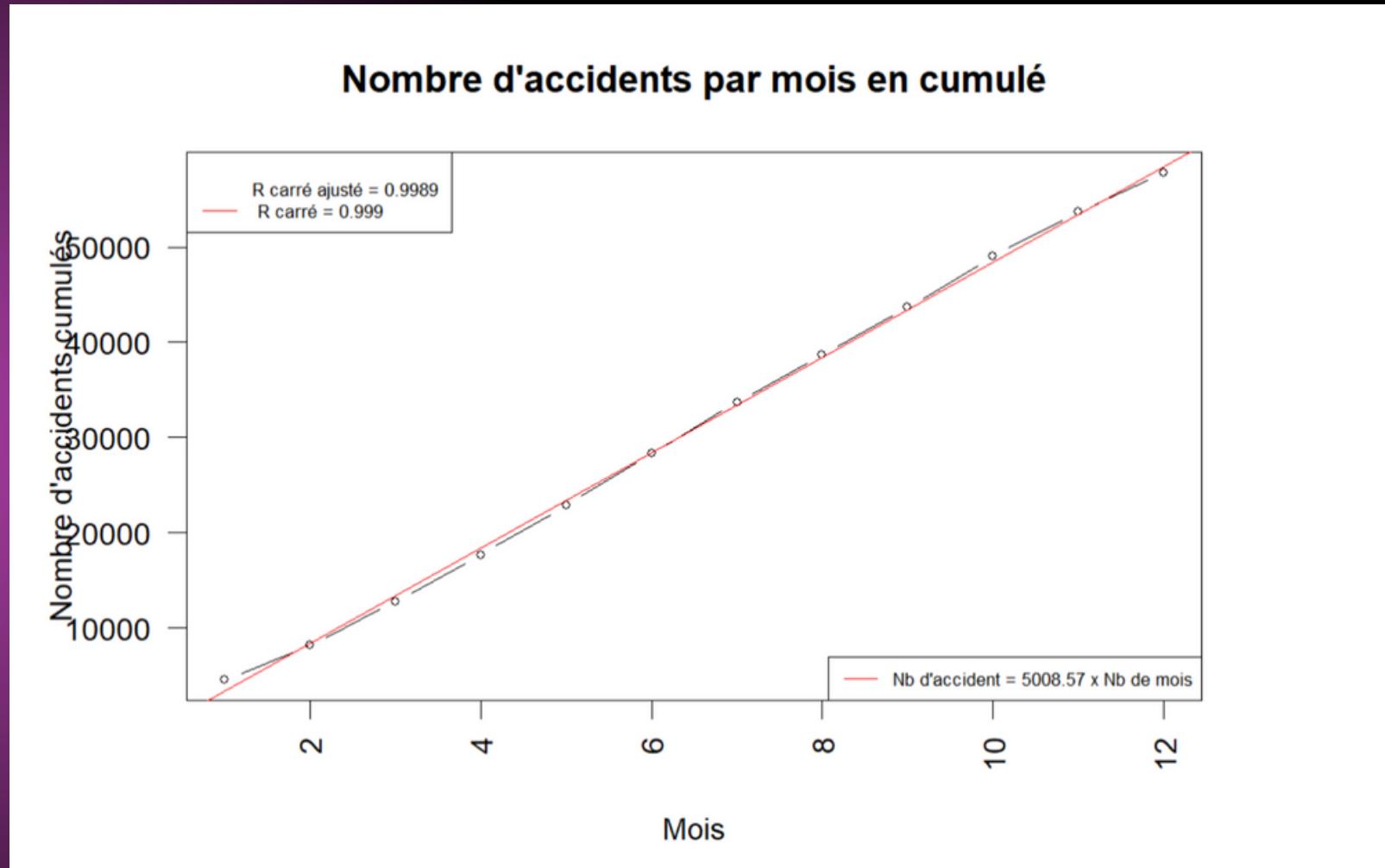
Taux
d'accidents
grave par région



Taux d'accidents
grave par
département



Régression linéaire



INTERVAL DE CONFIANCE

	2.5 %	97.5 %
(INTERCEPT)	-1727.295	-1175.456
SEMAINE	1143.870	1161.990

Diagrammes en mosaïque

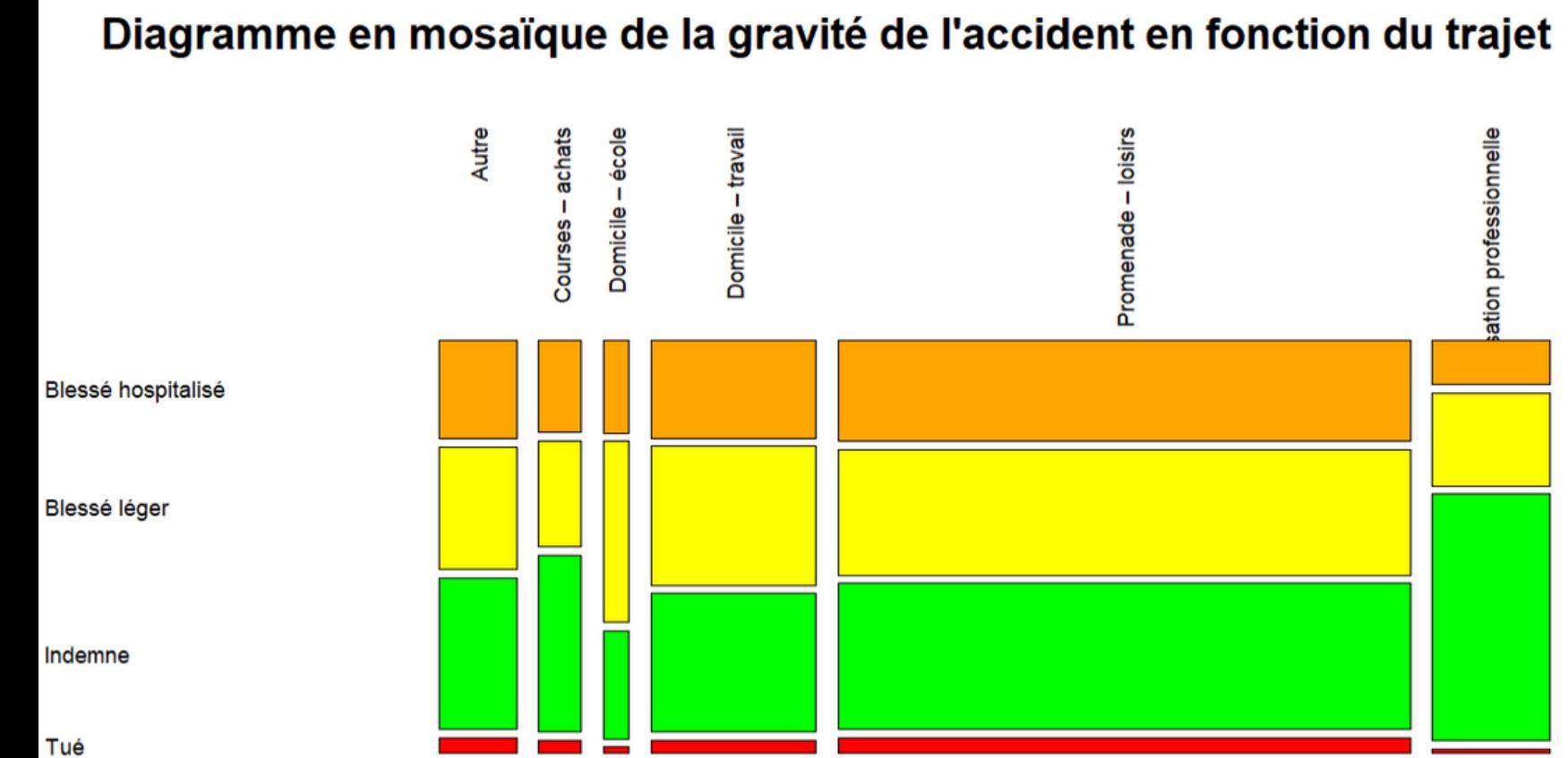
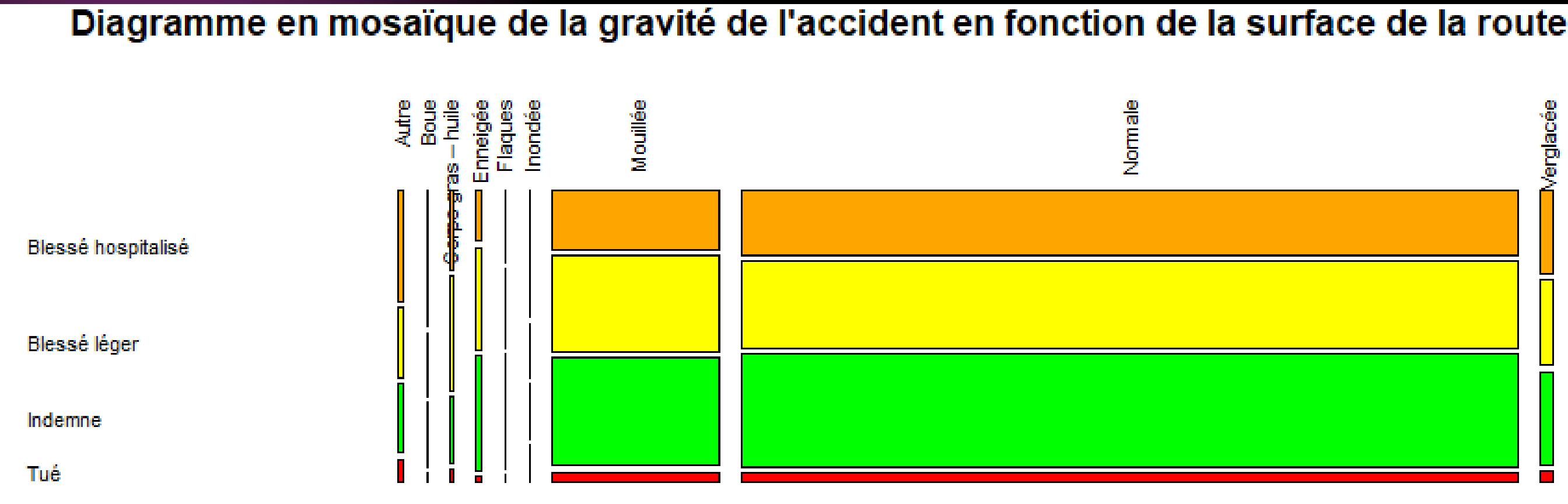
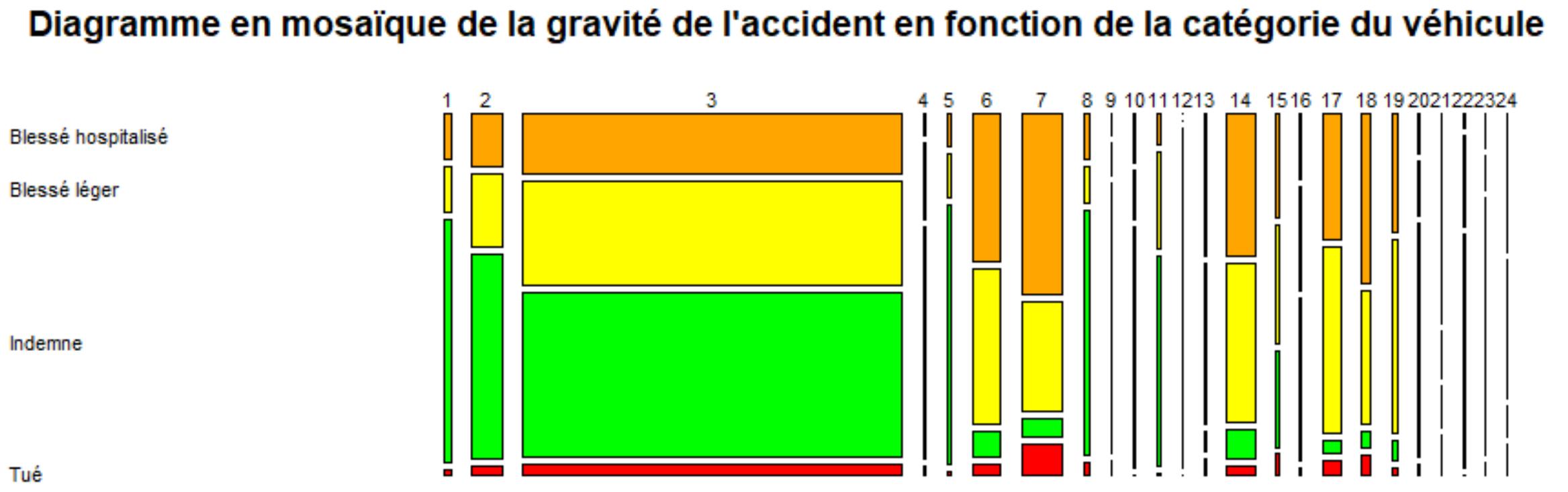


Diagramme en mosaïque de la gravité de l'accident en fonction de la surface de la route



Diagrammes en mosaïque

1 = PL seul > 7,5T
2 = VU seul 1,5T <= PTAC <= 3,5T avec ou sans remorque
3 = VL seul
4 = Autocar
5 = PL > 3,5T + remorque
6 = Cyclomoteur <50cm³
7 = Motocyclette > 125 cm³
8 = Tracteur routier + semi-remorque
9 = Tracteur agricole
10 = PL seul 3,5T <PTCA <= 7,5T
11 = Autobus
12 = Train
13 = Scooter > 125 cm³
14 = Scooter < 50 cm³
15 = Voiturette (Quadricycle à moteur carrossé) (anciennement "voiturette ou tricycle à moteur")
16 = Autre véhicule
17 = Bicyclette
18 = Motocyclette > 50 cm³ et <= 125 cm³
19 = Scooter > 50 cm³ et <= 125 cm³
20 = Engin spécial
21 = Quad lourd > 50 cm³ (Quadricycle à moteur non carrossé)
22 = Tramway
23 = Tracteur routier seul
24 = Quad léger <= 50 cm³ (Quadricycle à moteur non carrossé)



Conclusion

- Importance du nettoyage des données
- exploité par les compagnies d'assurances afin de fixer les prix selon les critères les plus accidentogènes