



le cnam
Ile-de-France

le cnam

- Conservatoire National des Arts et Métiers -

- Master 1 Droit, économie et gestion, mention Sciences criminelles et criminologie -

CRM 205 Territoires et Cartographie :

« Morts violentes en famille »

**Étude des homicides intrafamiliaux basée sur des articles de journaux
de 2011 à 2015**

LEFEBVRE Stéphanie - ROBERT Olivier - ROGER Olivier

Directeur du Master :

Professeur Alain BAUER

Responsable de l'UE :

Professeur Michel BERA

SOMMAIRE

INTRODUCTION

I. DÉFINITION DU SUJET

- 1) Définition issue de la psychiatrie
- 2) Définition journalistique
- 3) Définition juridique : qualification pénale
- 4) Définition de notre sujet d'étude

II. RECUEIL ET PRÉSENTATION DES DONNÉES

- 1) Présentation des données nécessaires à notre projet
- 2) Méthodologie de collecte des données
- 3) Présentation de la base données créée

III. ANALYSE STATISTIQUE DESCRIPTIVE DU PHÉNOMÈNE

- 1) Avant-propos concernant le logiciel R
- 2) Les homicides en France entre 2007 et 2016
- 3) Analyse du phénomène de « drames familiaux »
- 4) Analyse concernant la victime du drame familial
- 5) Analyse concernant l'auteur du drame familial

IV. ANALYSE CARTOGRAPHIQUE

- 1) Répartition des « drames familiaux » sur le territoire français
- 2) Hypothèse 1 : Corrélation spatiale nombre d'armes / nombre d'homicides
- 3) Normalisation du nombre de victimes par département
- 4) Hypothèse 2 : Influence des saisons
- 5) Hypothèse 3 : Influence du taux de chômage «environnemental »
- 6) Hypothèse 4 : Taux de pauvreté

V. CONCLUSION ET LIMITES DE LA DÉMARCHE

ANNEXES

INTRODUCTION :

La cartographie est basée sur la géographie, « science qui a pour objet l'étude des phénomènes physique, biologiques, humains localisés à la surface du globe terrestre et spécialement l'étude de leur répartition, des forces qui les gouvernent et de leurs relations réciproques »¹

« La statistique est l'ensemble des méthodes visant à décrire, résumer et interpréter des phénomènes dont le caractère essentiel est la variabilité. L'analyse s'effectue sur l'ensemble des données résultant de l'observation préalable de ces phénomènes. Le paradigme de la statistique repose sur l'idée qu'un ensemble d'individus forment une population et que c'est cette population, et non l'individu, qui est l'objet de l'étude. »²

La criminologie travaille sur des phénomènes de faible ampleur. Il est difficile, mais pas impossible, de construire des théories et d'expliquer des phénomènes grâce aux outils statistiques. En ce sens, la criminologie se rapproche fortement de la « statistique épidémiologique »³. La statistique épidémiologique peut avoir 4 facettes :

- épidémiologie descriptive : évaluer la fréquence d'une maladie ;
- épidémiologie analytique : comparer cette fréquence entre des groupes d'individus dans des circonstances diverses, afin de mettre en évidence des facteurs de risque ;
- épidémiologie interventionniste : évaluer l'efficacité des interventions médicales et dont les essais thérapeutiques sont un cas particulier ;
- épidémiologie prédictives : étude des facteurs pronostiques et à l'évaluation des moyens de diagnostics.

Appliquées à la criminologie, cela donnerait :

- criminologie descriptive : évaluer la fréquence d'un crime ou d'un délit
- criminologie analytique : comparer cette fréquences entre des groupes d'individus par exemple : récidiviste/non-récidiviste, condamnés à une peine de prison/condamnés à une peine alternative, etc. dans le but de mettre en évidence des facteurs de risque ;
- criminologie évaluer l'efficacité de la réponse pénale et de la chaîne pénale (police, justice, pénitentiaire) ;

1 « Géographie et statistique » Vigneron E., Collection Que sais-je ?, aux Éditions Presses Universitaires de France

2 « La statistique » Vessereau A. Collection Que sais-je ?, aux Éditions Presses Universitaires de France

3 « Statistique épidémiologique » Renaud A. Collection Que sais-je ?, aux Editions Presses Universitaires de France

- criminologie prédictive : développement d'une politique de prévention criminelle, voir de prédiction.

Le langage R est avant tout un langage informatique dédié aux statistiques et à la science des données⁴. Pour pouvoir l'utiliser, il faut installer **LE LOGICIEL GNU R. RSTUDIO** est un environnement de développement plus facile d'utilisation. Rstudio propose notamment une écriture semi-automatique des fonctions, signale les erreurs de syntaxes, permet de visualiser en permanence les variables utilisées, etc.... Vu les nombreux avantages de Rstudio, c'est avec cette interface que nous travaillerons pour ce projet.

Ce projet sera l'occasion

- de définir un sujet d'étude criminologique
- d'émettre des hypothèses d'études concernant un phénomène criminologique
- de découvrir le langage R et ses utilisations possibles
- d'établir une base de données fondée sur la collecte de faits caractérisant le sujet choisi
- d'effectuer une étude statistique à partir de notre base de données
- de cartographier ce phénomène par rapport à nos hypothèses

I. Définition du sujet

1) Définition issue de la psychiatrie

« Les crimes intrafamiliaux sont commis par une personne contre un membre d'une même famille, lié par le sang, par alliance ou par adoption. La criminalité intrafamiliale est assez rare mais crée toujours des situations qui débordent largement le strict cadre familial et touchent l'ensemble de la société. »⁵

Cette définition permet de définir clairement le lien indispensable existant entre la victime et l'auteur : ils doivent être lié par le sang, l'alliance ou l'adoption.

⁴ [https://fr.wikipedia.org/wiki/R_\(langage\)](https://fr.wikipedia.org/wiki/R_(langage))

⁵ Bram Khemiri, Nesrine, Imene Ghazali, et Rym Ridha. « Schizophrénie et crimes intrafamiliaux. Étude analytique et comparative à propos de 80 cas », *L'information psychiatrique*, vol. volume 88, no. 10, 2012, pp. 839-846.

2) Définition journalistique

L'expression « drame familial » est fréquemment utilisée par les journalistes car elle suscite une vive émotion au sein de la société. Pourtant, elle diffère de la définition issue de la psychiatrie. Le lien reste le même, mais le crime est restreint aux homicides⁶. Ainsi, les viols peuvent être des crimes intrafamiliaux, mais qui ne seront pas pris en compte dans cette étude.

3) Définition juridique : qualification pénale

Le meurtre est le fait de donner volontairement la mort à autrui (art. 221-1 du Code pénal).

L'assassinat est un meurtre commis avec préméditation ou guet-apens (art. 221-3 du Code pénal).

Les violences volontaires ayant entraîné la mort sans intention de la donner sont prévues par l'art 222-7 du Code pénal.

4) Définition de notre sujet d'étude

Au vue de ces éléments, nous avons décidé que le phénomène étudié concernera les homicides, où l'auteur a un lien de parenté, par alliance, par le sang, ou par l'adoption, avec sa victime.

II. Recueil et présentation des données

1) Présentation des données nécessaires à notre projet

L'article 15 de la Déclaration des droits de l'homme et du citoyen de 1789 dispose que "la société a le droit de demander compte à tout agent public de son administration"⁷. Pour cette raison, les données administratives produites par l'État doivent être transparentes et accessibles à tous les citoyens⁸. Ce partage et cette mise à disposition des données constituent « l'open data ». Ainsi, en 2014, la France était, selon l'ONU, le 4^e pays au monde en matière d'administration numérique et le 1^{er} Européen. Ces données émanant de l'État, elles sont considérées comme fiables.

En ce qui concerne les données géographiques, les sources sont nombreuses. Nous citerons les plus connues et les plus complètes :

- GoogleMap ;
- OpenStreetMap : projet international, débuté en 2004, dans le but de créer un carte libre du monde⁹. La carte est libre car elle est modifiable à tout moment et par chaque internaute.

6 Article 221-1 du Code Pénal : « Le fait de donner volontairement la mort à autrui constitue un meurtre. »

7 <http://www.gouvernement.fr/action/l-ouverture-des-donnees-publiques>

8 Loi CADA 1978

9 <http://openstreetmap.fr/projet>

Elle comprend les routes, voies ferrées, rivières, forêts, bâtiments, etc. Elle est utilisable avec le packages « OpenStreetMap » ou « OSM » ;

- GEOFLA : est une base de données créée par l'Institut Géographique National.

Notre sujet n'ayant pas de bases de données pré-existantes, nous allons devoir en constituer une.

2) Méthodologie de collecte des données spécifiques au sujet traité

L'un des auditeurs de notre groupe est journaliste. Dans le passé, il a couvert des drames familiaux. Nous sommes donc en possession d'une première série de données utilisables et très précises

Un autre auditeur a découvert des sites recensant de très nombreux articles de journaux. Le blog principal (<https://aucoeurdudrame.blogspot.fr>) débute en Août 2011. Nous utiliserons les catégories « drame familial », dans laquelle figure 498 articles, et « drame passionnel » regroupant 14 articles.

Pour compléter cette série, nous avons effectué une recherche « drame familial » sur le site internet de France 3 région (<http://france3-regions.francetvinfo.fr/recherche?s=drame+familial>), puis une seconde recherche sur Google (<https://www.google.fr>) ainsi que sur d'autres blogs (<http://blogs.nouvelobs.com/explore/posts/tag/drame+familial?&from=255>) .

Cette méthode nous permet de confronter et recouper plusieurs bases de données pré-existantes. Nous ne pensons pas que notre étude soit exhaustive mais elle tend à être la plus complète possible. Nous avons retenu que les assassinats, les homicides volontaires ou violences suivies de mort commis à l'encontre de sa famille, homme ou femme.

Afin de recueillir suffisamment de données ces recherches ont été faites sur plusieurs années que nous nous sommes réparti comme tel :

- auditeur 1 : années 2010 et 2014
- auditeur 2 : années 2011 et création d'une base regroupant tous les faits par départements
- auditeur 3 : années 2012 et 2015

Constituer cette base de données sera l'étape la plus délicate et fastidieuse de ce projet. Nous avons pour des raisons pratiques traité que la France métropolitaine.

3) Présentation de la base de données créée

La base de données est créée à l'aide du logiciel Excel et sera enregistrée au format .CSV pour être facilement exploitable avec le logiciel R. En voici un aperçu :

Drames, Familiaux, Projets, Bera, ods - Excel

stéphanie

Fichier Accueil Insérer Dessin Mise en page Formules Données Révision Affichage Q Dites-nous ce que vous voulez faire

Calibri 11 A A

Renvoyer à la ligne automatiquement

Personnalisée

0 % 000 000

Mise en forme conditionnelle

Mettre sous forme de tableau

Styles de cellules

Insérer

Supprimer

Format

Cellules

Trier et Rechercher

Filter & sélectionner

Édition

Coller

Police

Alignement

Nombre

Presses-papiers

A1

moment_journee

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
	moment_journee	moment_semaine	mois	annee	ville	code_postal	departement	lien_parente	age_auteur	sexe_auteur	metier_auteur	nombre_victime	age_victime	sexe_victime	mode_operation	suite	Q
2	nuit	Week-end	janvier	2011	montereau	89470	yonne	En-compagn	34	homme	na	1	28	femme	couteau	interpelle	D100RUn+ter
3	na	semaine	janvier	2011	Pont-de-metz	60490	somme	mar	60	homme	retraite	3	30	homme	couteau	suscite	gemaire-end
4													30	femme	couteau		
5													20	femme	couteau		
6	midi	semaine	janvier	2011	boson	06830	haut-rhin	amant	27	homme	na	1	20	femme	brûle	interpelle	coise-brûler
7	nuit	na	janvier	2011	boulancourt	77760	seine et marne	pere	41	homme	maçon	1	6	filie	etouffe	interpelle	r-avec-surst
8	mid	semaine	janvier	2011	quimper	29000	bretagne	Belle seue	23	femme	chomage	1	24	femme	couteau	interpelle	http://www.d
9	na	Week-end	janvier	2011	rouvenzhelm	67480	alsace	mar	45	homme	boucher	1	38	femme	strangulation	interpelle	http://www.d
10	soir	semaine	mars	2011	epenay	51200	marne	mere	35	femme	na	1	8	garçon	etouffe	interpelle	http://www.le
11	matin	semaine	mars	2011	compigne	60200	picardie	En-compagn	41	homme	na	1	45	femme	arme à feu	interpelle	naïl-avert-pa
12	na	semaine	avril	2011	cretel	94000	val de marne	En-compagn	47	homme	na	3	45	femme	couteau	interpelle	https://www.s
13													15	garçon	couteau		
14													45	garçon	couteau		
15	mid	semaine	avril	2011	Heiltz-le-Maurup	51340	marne	germe	53	homme	retraite	1	84	femme	arme à feu	interpelle	RTFQ04080
16	nuit	na	avril	2011	nantes	44000	loire atlantique	pere	50	homme	commercial	5	48	femme	arme à feu	hute	pres-lafaire

Notre base de données se veut ambitieuse dans le nombre de critères recueillis pour chaque « drame familial ».

Toutefois, nous devons admettre certaines limites :

- le métier de l’auteur des faits est important mais malheureusement, les articles de journaux ne précisent que très rarement cette information ;
- la découverte des corps et la rédaction de l’article peuvent être fait quelques heures après le drame ou plusieurs jours après, ce qui rend difficile la détermination du moment de la commission de l’homicide ;

Enfin, Monsieur Béra et Monsieur Langlade, nous ont tous deux précisé que pour que notre étude soit pertinente, il nous faudrait environ 500 cas. Ce chiffre semble difficilement atteignable par manque de temps. Nous espérons pouvoir mener notre études sur 400 cas, répartis sur 5 ans.

VI. Analyse statistique du phénomène

Notre sujet traitant d'un phénomène peu étudié et reposant sur une base de données créée de toute pièce, la première étape sera de présenter nos résultats sous forme de graphiques.

Après une courte présentation des homicides en France, nous avons choisi de :

- Dresser un « portrait robot » de l'auteur : âge, sexe, mode opératoire
- Déterminer le profil des victimes afin de déterminer quel type de personne est en position de vulnérabilité face à ces drames familiaux : enfants ou adultes ? Homme ou femme ? Etc...

1) Avant propos concernant le logiciel R

Le logiciel R fonctionne grâce à des packages. Ceux-ci sont stockés sur des serveurs universitaires. La première étape est donc d'exécuter la commande « `install.packages (" Nom_du_package ")` », un dossier, nommé « `Nom_du_package` » se crée dans le répertoire « `C:\Users\Documents\R\win-library\3.3` ». Ce dossier contient toutes les fonctions en rapport avec le package.

La commande « `install.package()` » ne doit être exécuter qu'une seule fois.

Une fois le package installé, il est nécessaire d'appeler les fonctions existantes dans le package. Cela est possible grâce à la commande « `Library ()` » et doit être exécutée à chaque fois que vous lancez R.

Chaque librairie comprend des fonctions utilisant des paramètres, placés entre parenthèses, après l'appel de la fonction. Pour connaître le fonctionnement de chaque fonction, il est nécessaire de se référer à la documentation du package, au format PDF, accessible sur internet. Via Rstudio, il est également possible de rechercher des informations dans l'onglet « Help » de la fenêtre située en bas à droite.

Exemple :

```
install.packages ( " ggplot2 " )
```

```
library ( " ggplot2 " )
```

Doc PDF : <https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf>

fonctions : `geom_bar`, `aes`, `autoplot`, `geom_label`, etc.....

paramètres : `geom_bar (mapping = NULL, data = NULL, stat = "count", position = "stack", width = NULL, binwidth = NULL, na.rm= FALSE, show.legend = NA, inherit.aes = TRUE)`

2) Les homicides en France entre 2007 et 2016

Le fichier « État 4001 » comprend les crimes et délits portés à la connaissance de la gendarmerie et de la police. Il ne comporte que les faits faisant l'objet d'une procédure judiciaire transmise au parquet, à la suite d'une plainte ou d'une enquête de police pour des faits graves¹⁰.

10 <https://sig.ville.gouv.fr/page/112> – Système d'Information Géographique de la politique de la ville

Les homicides étant des faits graves, sanctionnés par l'article 221-1 du Code Pénal, ils figurent dans le fichier « État 4001 ». L'homicide est le crime le mieux recensé dans les pays développés comme la France et représente un indicateur du degré de violence d'une société.

L'État 4001 est constitué de 107 index qui, selon la classification de l'Observatoire national de la délinquance et des réponses pénales (ONDRP), peuvent faire l'objet de cinq grands regroupements :

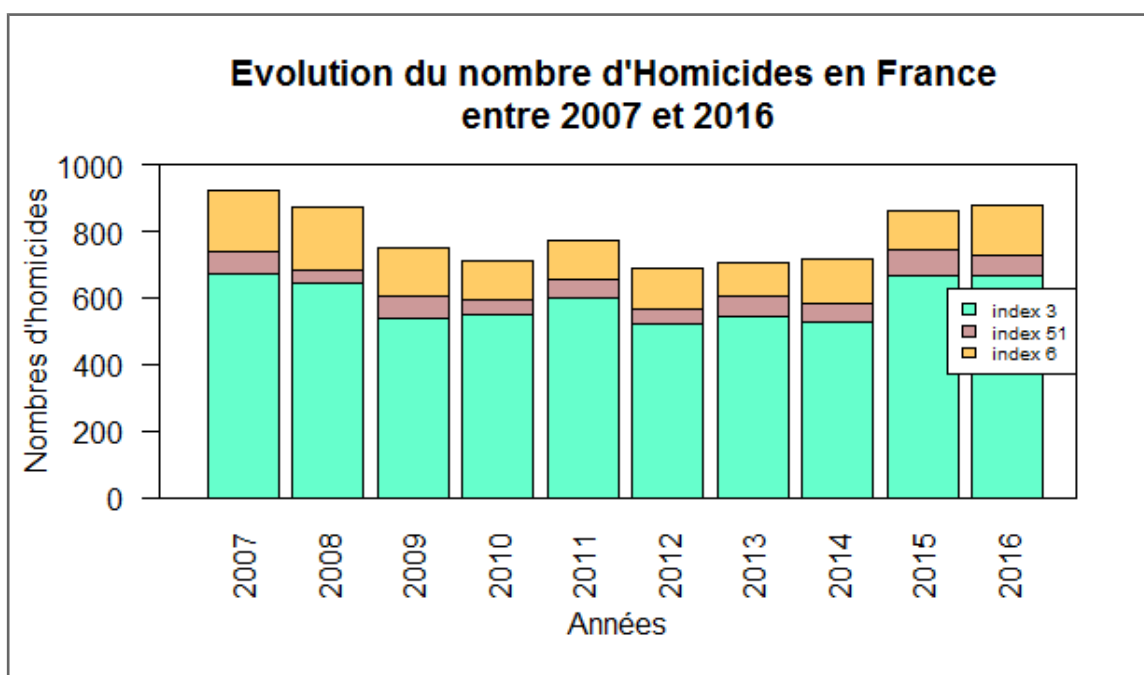
- atteintes aux biens
- atteintes aux personnes (ou atteintes volontaires à l'intégrité physique)
- infractions révélées par l'action des services
- escroquerie et infractions économiques et financières
- autres infractions.

Dans le « guide de méthodologie statistique », datant de 2015, transmis par Monsieur Besson, nous comprenons de quelle manière sont indexés les crimes et délits (Cf Annexe 1 : Indexation des infractions). Les homicides entrent dans la catégorie « atteintes aux personnes » et sont répartis dans plusieurs catégories.

Pour notre cas d'étude, nous ne retiendrons que 3 catégories :

- homicide pour d'autres motifs : **Index 3** : ne comprend pas les homicides issus de règlements de compte ou à l'occasion de vol. Les victimes ont obligatoirement plus de 15 ans ;
- homicide commis contre enfants de moins de 15 ans : **Index 51** : complète l'index 3 en comptabilisant les victimes de moins de 15 ans ;
- coups et blessures volontaires suivis de mort : **Index 6** : les coups ont entraîné la mort sans intention de la donner, quel que soit l'âge de la victime

L'État 4001 fait partie des données administratives accessibles aux citoyens sur le site data.gouv.fr, nous avons pu le télécharger au format .XLSX que nous avons converti en .CSV afin de l'utiliser plus aisément sous R.



A partir de ce fichier, nous avons construit, sous R, un histogramme cumulé nous permettant de visualiser l'évolution du nombre d'homicides sur 10 ans, grâce à la fonction « `barplot()` ».

Le rapport « Insécurité et délinquance en 2016 », publié par Interstats¹¹, nous rappelle que si le nombre d'homicide a augmenté ces deux dernières années, c'est à cause des attentats. Il nous révèle que:

- 39 % des victimes sont des femmes,
- les classes d'âge les plus touchées sont les 15-29 ans et les 30-44 ans,
- presque un homicide sur 5 a lieu dans la sphère familiale,
- 55 % des homicides de mineurs de 15 ans a lieu dans la sphère familiale.

Ces informations seront importantes pour relativiser et analyser notre base de données.

¹¹ <http://www.interieur.gouv.fr/Interstats/Actualites>

```

# ----- CODE HISTOGRAMME HOMICIDE ETAT 4001 -----
# Importer le fichier .CSV état 4001 homicide de 1996 à 2016
hom <- read.csv2("homicide4001.csv", header=TRUE, sep="," , dec=".", quote="\"",
fill=TRUE)

# Création d'un fichier data frame
data.frame(hom)

# Création d'une matrice avec les années nécessaires pour définir l'abscisse
annee = c(2007,2008,2009,2010,2011,2012,2013,2014,2015,2016)

# Création d'une matrice pour chaque type d'infraction (type d'homicide)
homicide_autre = c(671,645,540,548,601,521,546,526,669,669)
homicide_mineur = c(69,39,65,45,55,46,62,56,75,57)
coup_mort = c(182,191,144,121,115,123,96,135,120,154)

# Avant affichage, regrouper les données sous forme de barres cumulées
cumul_homicide <- cbind (homicide_autre,
                        homicide_mineur,
                        coup_mort )

# Affichage des histogrammes cumulés grâce à la fonction barplot()
Barplot (t (cumul_homicide),                                # Données
        beside= F,                                           # Position des barres Beside = F ou T
        col= c ("#66FFCC","#CC9999","#FFCC66"),             # couleurs des barres
        xlab= "Années",                                       # Label (nom) en abscisse
        ylab= "Nombres d'homicides",                         # Label en ordonnée
        names=annee,
        main = "Evolution du nombre d'Homicides en France \nentre 2007 et 2016", # Titre
        las = 2,
        horiz = F,                                           # barres horizontales ou verticales horiz = F ou T
        xlim = c(0,12),                                     # valeurs mini et maxi en abscisse
        ylim = c(0,1000)                                    # Valeurs mini en maxi en ordonnée
)

# Légende
legend(x="right",
      legend=c ("index 3",  #Nombre d'homicides pour un autre motif",
                "index 51", #Nombre d'homicides de mineurs de 15ans",
                "index 6"), #Coups et blessures ayant entraînés le décès"
        Cex=0.6,  # Taille de la police
        bg = '#FFFFFF',  # Couleur de fond
        fill=c("#66FFCC","#CC9999","#FFCC66")) # Couleur de la légende

# Fonction permettant d'encadrer le diagramme
Box()

```

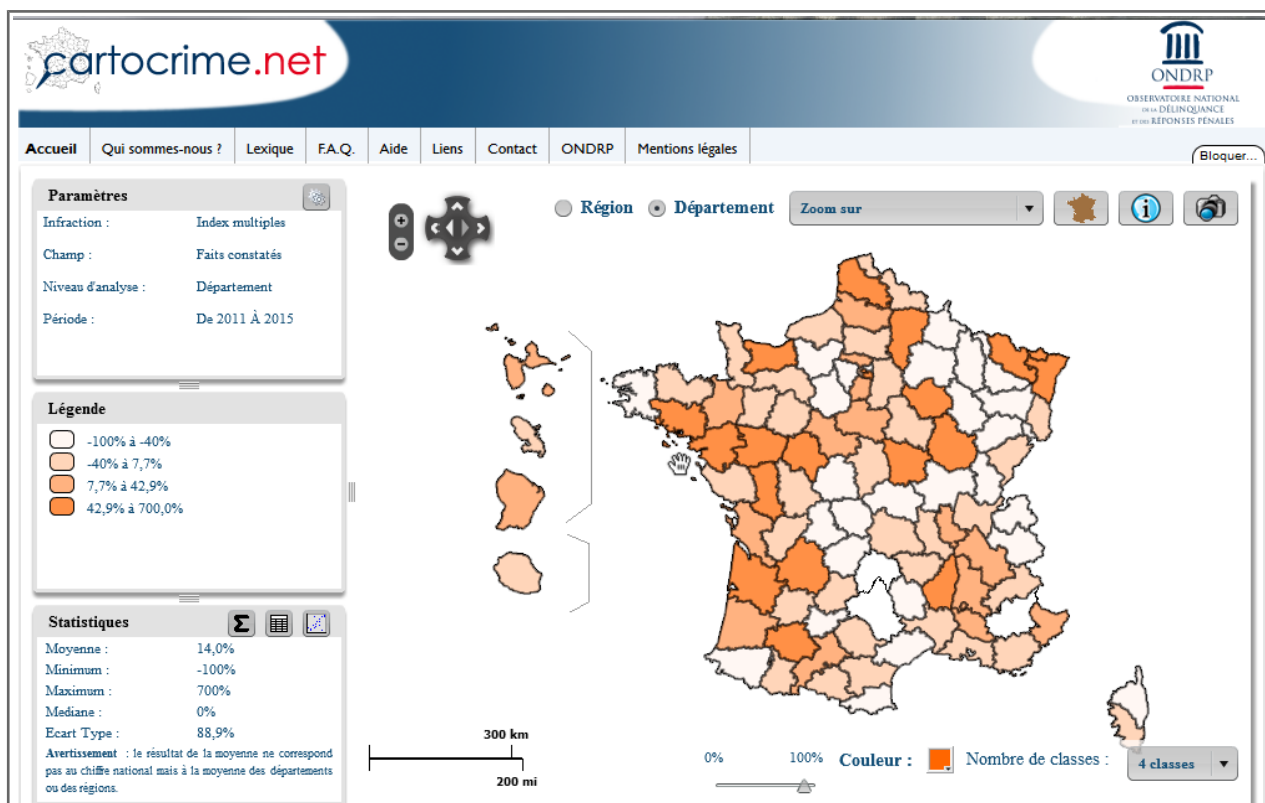
L'Observatoire National de la Délinquance et de la Réponse Pénale (ONDRP) a développé un outil qui permet de visualiser les crimes sous forme de carte. Outil est disponible sur le site CartoCrime.net.

Sur ce site nous avons la possibilité de voir la carte des crimes et délits que nous sélectionnons. La catégorisation des crimes et délits respectent les index de l'État 4001. Il nous est également possible de choisir la période qui nous intéresse.

En sélectionnant les index suivants :

- homicide pour d'autres motifs : Index 3 ;
- homicide commis contre enfants de moins de 15 ans : Index 51 ;
- coups et blessures volontaires suivi de mort : Index 6

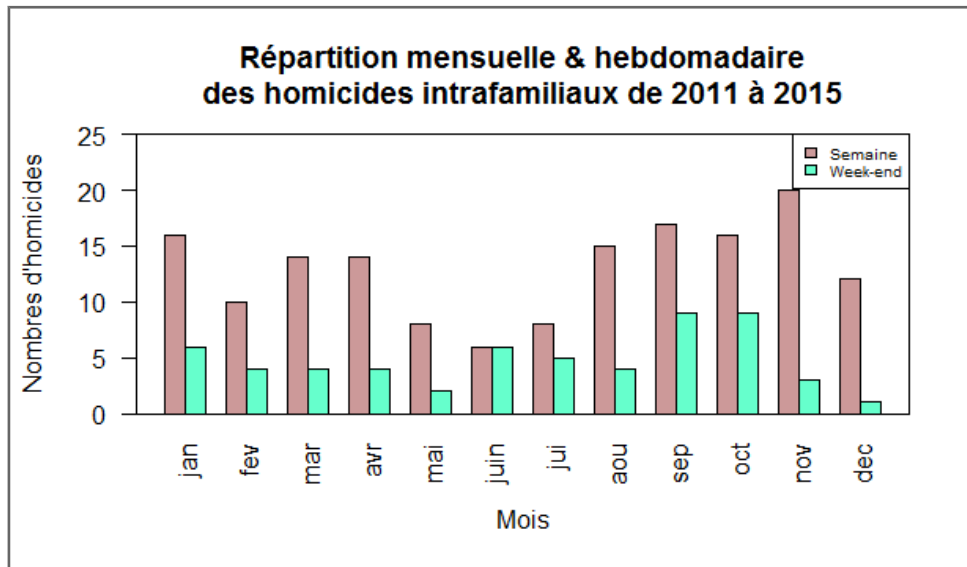
entre 2011 et 2015, nous avons obtenu la carte suivante :



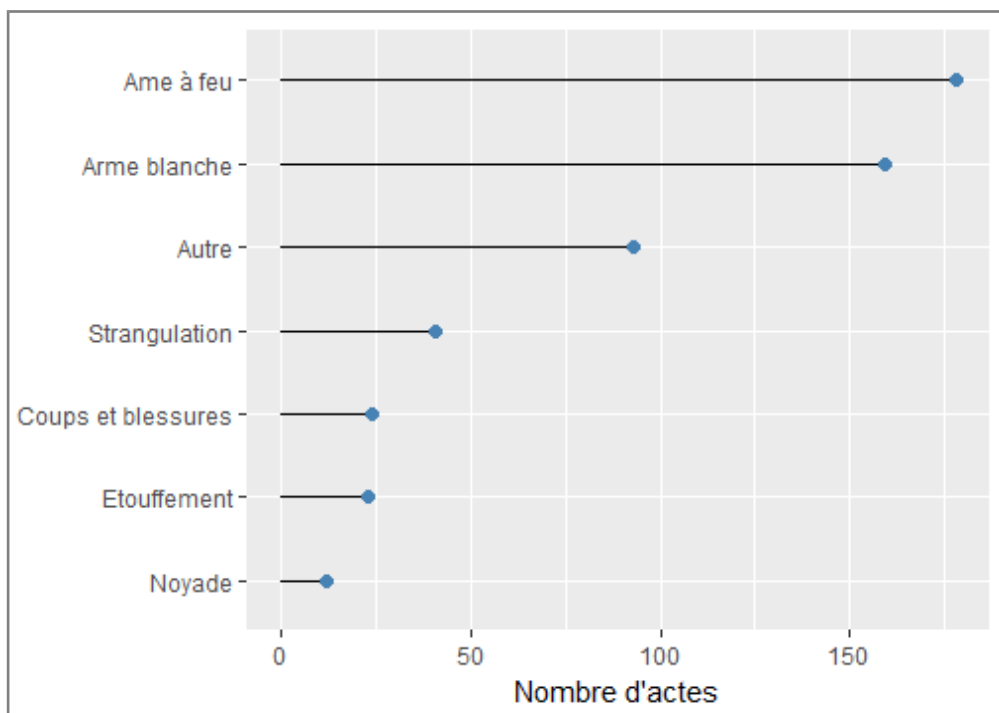
Cette carte pourra nous servir de référence, dans une étude comparative, pour répondre à la question suivante : « Les homicides intrafamiliaux se produisent-ils sur les territoires où le nombre d'homicide est le plus élevé ? »

3) Analyse du phénomène de « drames familiaux »

Notre étude porte sur 439 drames familiaux ayant fait 566 victimes.



Avant de commencer cette étude, nous avions en tête que les homicides intrafamiliaux se produisaient majoritairement le week-end car c'est à cette période que la famille se trouve réunie. Cet histogramme montre que notre a priori était faux. Les drames se produisent plutôt en semaine et le nombre semble plus élevé sur la période allant de Août à décembre.



Ce graphique présente le nombre de drames familiaux – c’est-à -dire le nombre d’acte et non le nombre de victimes - par mode opératoire.

Nous constatons que les armes les plus fréquemment utilisées sont de loin les armes à feu et armes blanches. La catégorie « autre » regroupe des moyens très divers et très peu fréquents. On y trouve : l’utilisation « d’armes par destination » (arbalète, barre de fer, marteau, pilon à mortier, hache, hachoir, harpon, et masse) mais aussi d’autre modes opératoires tel que l’incendie et l’empoisonnement. Il aurait été intéressant de distinguer le mode opératoire en fonction du sexe de l’auteur mais nous le ne ferons pas par manque de temps.

```
# Diagramme en sucette
library(ggplot2)
library(ggalt)

# Nettoyage des données
drametotal[drametotal$mode_operatoire=="",]<-NA
drametotal<-drametotal[complete.cases(drametotal$mode_operatoire ),]

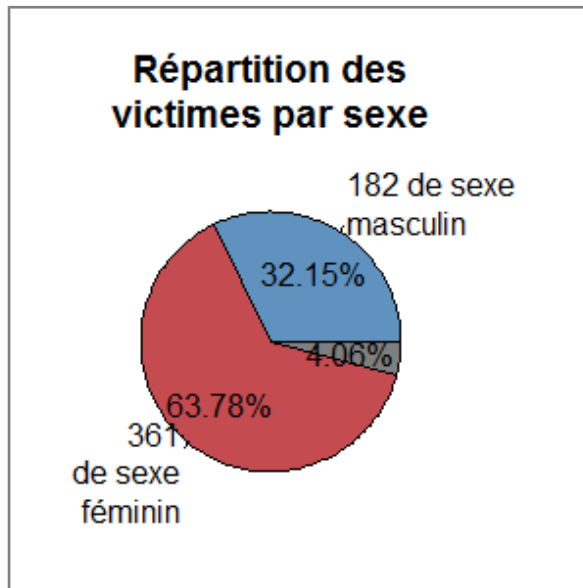
af = 0 ; ar = 0 ; st = 0 ; et = 0 ; autre = 0; no =0 ; cp =0;

n = nrow(drametotal)
for (i in 1:n)
{ if (drametotal$mode_operatoire[i] == "arme a feu"){ af = af +1}
  if (drametotal$mode_operatoire[i] == "couteau"){ ar = ar +1}
  if (drametotal$mode_operatoire[i] == "strangulation"){ st = st +1}
  if (drametotal$mode_operatoire[i] == "etouffement"){ et = et +1}
  if (drametotal$mode_operatoire[i] == "noyade"){ no = no +1}
  if (drametotal$mode_operatoire[i] == "coups"){ cp = cp +1}
  else { autre = autre +1 }
  i = + i
}

# Organisation des données
modeoperatoire <- read.csv(text=" ModeOperatoire,Frequence
    Ame à feu, af
    Arme blanche, ar
    Strangulation, st
    Noyade, no
    Coups et blessures, cp
    Etouffement, et
    Autre, autre
    stringsAsFactors=FALSE, sep=",", header=TRUE)

# Création du graphique à l'aide de GGLOT2
ggplot(modeoperatoire, aes(y= reorder(ModeOperatoire, Frequence), x =Frequence )) +
  geom_lollipop(point.colour="steelblue", point.size=2, horizontal=T) +
  ylab("Mode Opérateur") + xlab("Nombre d'actes")
+ labs(title="Répartition des homicides en fonction du mode opératoire",
  subtitle="Données entre 2011 et 2015",
  caption="SL / OR / OR") + coord_flip()
```

4) Analyse concernant la victime du drame familial



Ce graphique montre que les victimes sont très majoritairement de sexe féminin. Les 4 % dont nous ignorons le sexe correspond aux enfants, dont il n'est pas spécifié dans les articles de journaux si il s'agit d'une fille ou d'un garçon.

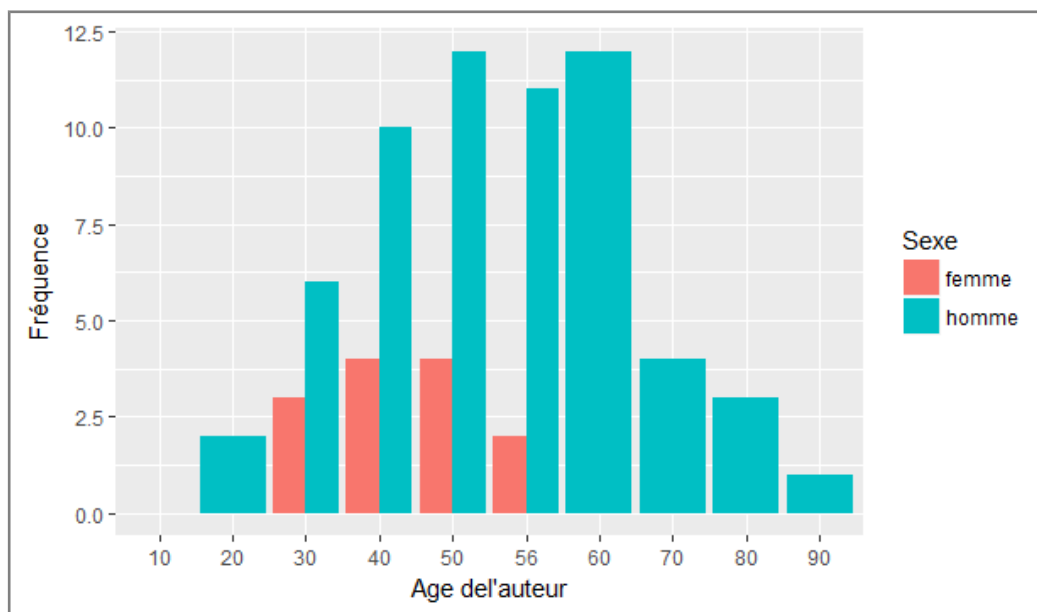
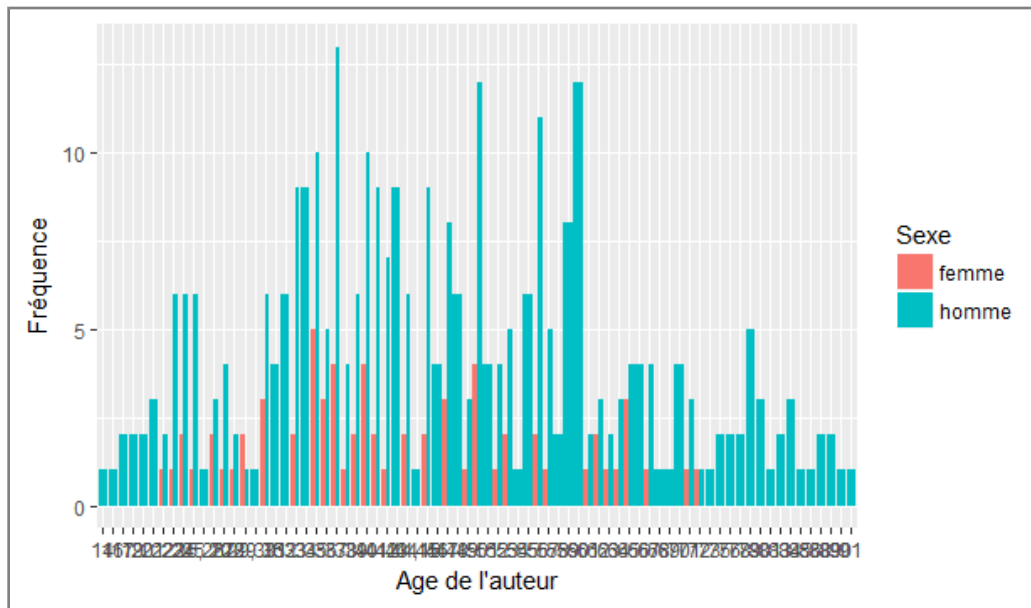
Il est possible de commencer par une étude statistique simple concernant l'âge de la victime :

- 173 mineurs de 15 ans, dont l'âge moyen est 5 ans
- 17 mineurs de 15 et plus
- 350 majeurs dont l'âge est connu et dont la moyenne est de 49 ans et demi

L'âge moyen d'une victime, toute catégorie confondue, est de 34 ans. On s'aperçoit que le nombre de mineurs, très jeune, fait très fortement baisser la moyenne.

La victime la plus jeune est un nourrisson de 7 jours tant dis que la plus âgée à 96 ans.

5) Analyse concernant l'auteur du drame familial



La répartition des auteurs suivant leur âge répond à une loi normale. Nous pouvons voir sur le second graphique que la répartition gaussienne des femmes est légèrement plus à gauche que celle des hommes.

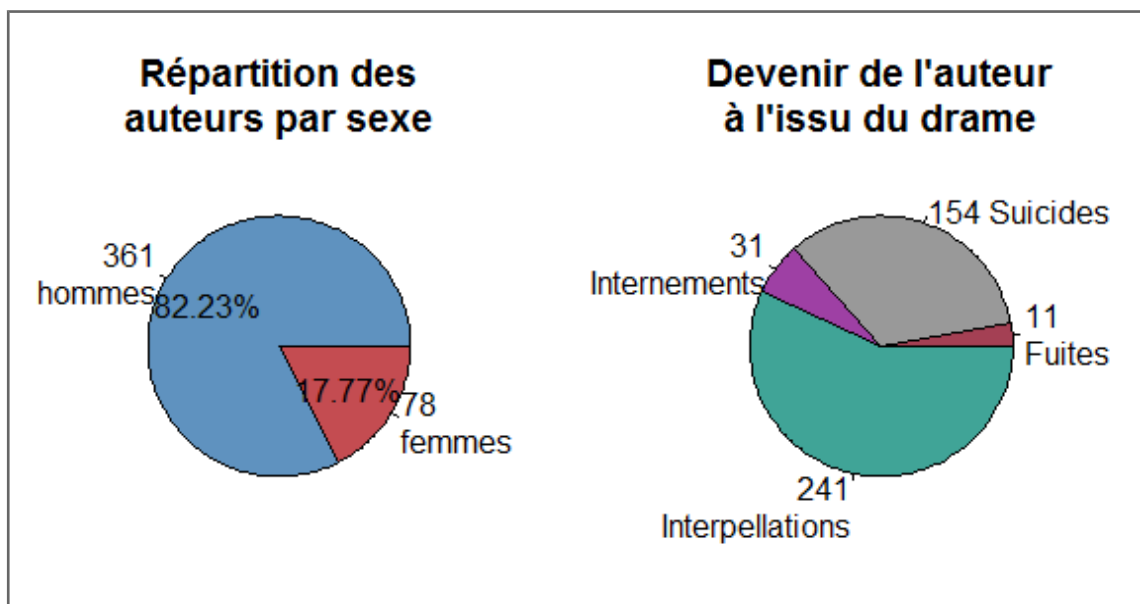
L'âge moyen de l'auteur est 45 ans et demi, le plus âgé à 90 ans tandis que le plus jeune est mineur (14ans). Si l'on tient compte du sexe, l'âge moyen pour la femme est de 41 ans tant dis que celui des hommes est de 47 ans.

D'autres analyses statistiques peuvent être réalisées grâce aux nombreuses fonctions de R.


```

#----- CALCULS STATISTIQUE -----
# Concaténation des données
drame <- rbind (drame2011, drame2012, drame2013, drame2014, drame2015)
# Calcul de la moyenne
mean ( drame$age_auteur )
# Affichage de la médiane
median ( drame$age_auteur )
# Calcul de la variance
var ( drame$age_auteur )
# Calcul de l'écart type
sd ( drame$age_auteur )

```



D'après les informations présentes dans les articles de presse, la très grande majorité des faits ont lieu au domicile du couple, durant une dispute conjugale. Les auteurs internés sont 22 hommes et 9 femmes, ayant commis des homicides principalement sur leur descendant ou leur ascendant.

Notre analyse statistique montre clairement que les auteurs sont majoritairement des hommes (plus de 82 %). Plus de la moitié sont interpellés directement après les faits, les quelques auteurs ayant pris la fuite, ne sont restés en cavale que quelques heures (ou ils ont été interpellés dans les 48h suivant le drame, ou ils se sont suicidés).

Le lien de parenté est présent dans notre base de donnée mais difficilement exploitable car un même auteur peut à la fois être père, fils et époux. Il devient difficile de catégoriser tout en ne comptant qu'un seul auteur.

```

# ----- CREATION DE DEUX CAMEMBERTS -----
# Diviser l'écran en deux
layout(matrix(1:2,1,2))

# Sexe de l'auteur
n = 0 ; n = nrow (drametotal) ; HomAut = 0 ; FemAut = 0

for (j in (1:n))
{ SexeAut = drametotal$sexe_auteur[j]
  if( SexeAut == "homme"){ HomAut = HomAut +1}
  if( SexeAut == "femme"){ FemAut = FemAut +1}
  j =+j}
sexe_auteurs <- c (HomAut, FemAut)

text_pie = function(vector,labels=c())
{ vector = vector/sum(vector)*2*pi
  temp = c() ; j = 0 ; l = 0
  for (i in 1:length(vector))
  { k = vector[i]/2
    j = j+l+k
    l = k
    text(cos(j)/2,sin(j)/2,labels[i]) }
  vector = temp}

PourcentageHomme <- ((HomAut *100)/(HomAut + FemAut))
PourcentageFemme <- ((FemAut *100)/(HomAut + FemAut))

couleurs = c("#5F92BF", "#C44C51")
pie (sexe_auteurs, labels = c("361\nhommes", "78\nfemmes"), main = "Répartition des\nauteurs par sexe",
     cex = 1, col = couleurs)
text_pie(sexe_auteurs, c("82.23%", "17.77%"))

# Diviser en type
n = 0 ; n = nrow (drametotal) ; enfuite = 0 ; suicide = 0 ; interpelation = 0 ; internement = 0 ; suite = 0

for (j in (1:n))
{ suite = drametotal$suite[j]
  if( suite == "fuite"){ enfuite = enfuite +1}
  if( suite == "suicide"){ suicide = suicide +1}
  if( suite == "interpelle"){ interpelation = interpelation +1}
  if( suite == "hospitalise"){ internement = internement +1}
  print (drametotal$suite[j])
  j =+j }
devenir_auteurs <- c (enfuite, suicide, internement, interpelation)

text_pie = function(vector,labels=c())
{ vector = vector/sum(vector)*2*pi
  temp = c()
  j = 0 ; l = 0
  for (i in 1:length(vector))
  { k = vector[i]/2
    j = j+l+k
    l = k
    text(cos(j)/2,sin(j)/2,labels[i])}
  vector = temp}

couleurs = c("#A44054", "grey60", "#9E40A4", "#40A497")
pie (devenir_auteurs, labels = c("11\nFuites", "154 Suicides", "31\nInternements", "\n241\nInterpellations"),
     main = "Devenir de l'auteur\nà l'issu du drame", cex = 1, col = couleurs)
text_pie(sexe_auteurs, c("82.23%", "17.77%"))

```

Les motivations de l'auteur sont étudiées dans la thèse « L'acte homicide-suicide : Revue de la littérature. A propos de 14 cas cliniques » de Anne-Sophie Chocard – 2002 – Université d'Angers – Faculté de médecine.

L. MILLET, P. VIVES et coll¹². dans leur article « Contribution à l'étude du suicide accompagné¹³ » de 1980, ont tenté d'analyser la signification psychopathologique des faits, à partir de 10 observations concernant des sujets ayant commis des « homicide-suicide ». Pour ces auteurs, la réalisation d'un suicide accompagné nécessite la rencontre de 3 facteurs associés :

1. l'existence d'un lien privilégié, mais de type immature, entre les personnes intéressées
2. une importante agressivité, associant le plus souvent auto- et hétéro-agressivité
3. des circonstances particulières favorisant le passage à l'acte.

Dans la suite de cette étude, nous chercherons à savoir si les circonstances particulières, présentées comme facteur favorisant le passage à l'acte, peuvent être environnementaux macroscopiques.

IV - ANALYSE CARTOGRAPHIQUE

La cartographie nous permet d'étudier la composition d'un territoire au travers de ses infrastructures, sa population, par tranche d'âge, revenus, situations familiales, son niveau d'instruction etc...

Ces informations sont disponibles sur 3 sites internet :

- <http://www.data.gouv.fr/fr/> : site mis en ligne par le gouvernement dont le but est de « partager, améliorer et réutiliser les données publiques » ;
- <https://www.insee.fr> : Institut National de la Statistique et des Études Économiques met à disposition des jeux de données statistiques classés par Thèmes / Niveau géographique / Catégories / Collections ;
- <http://ec.europa.eu/eurostat/fr/data/database> : Eurostat est l'office de statistique de l'Union européenne. La mission d'Eurostat est de fournir des statistiques de haute qualité pour l'Europe.

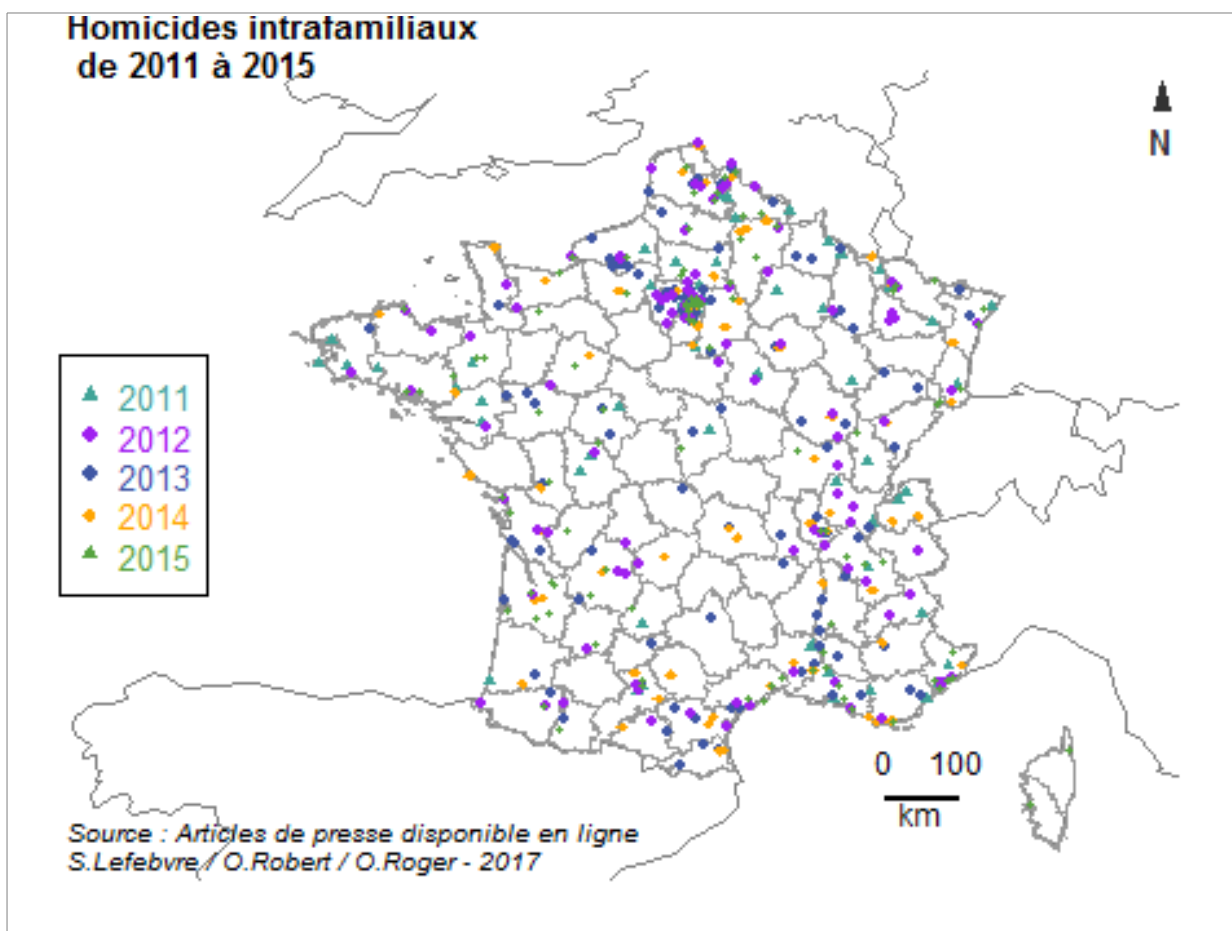
12 Millet L., Vives P., Moron P., Duguet AM., Arbus L. Contribution à l'étude du suicide accompagné. Médecine Légale, Toxicologie, 23, n°2, 1980 : 123-134.

13 l'expression « suicide accompagné » est une autre dénomination de l'expression « homicide-suicide »

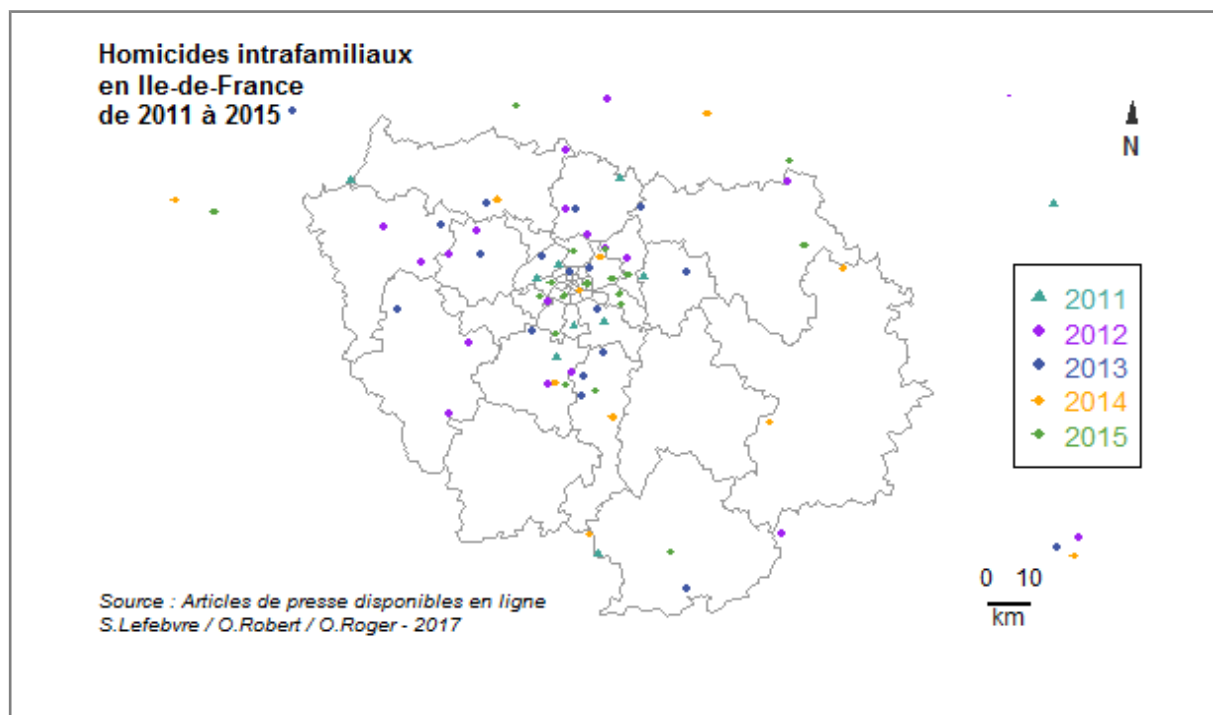
Dans cette partie, nous allons émettre des hypothèses que nous vérifierons à l'aide de la cartographie. A chaque hypothèse, nous associerons un fichier de données trouvé sur les sites data.gouv.fr, insee.fr ou ec.europa.eu/eurostat.fr.

Ces 3 sites présentent des données considérées comme fiables car ils mettent à disposition des données collectées par l'État français et l'Europe.

1) Répartition des « drames familiaux » sur le territoire français



Cette première carte représente la représentation des drames intrafamiliaux sur l'ensemble de la métropole. Elle n'est pas significative et nous ne pouvons pas en tirer d'analyse pertinente. Nous pouvons seulement constater que les homicides sont beaucoup plus nombreux en région parisienne que sur le reste de la France et que certains départements semblent épargnés par ce type de crime.



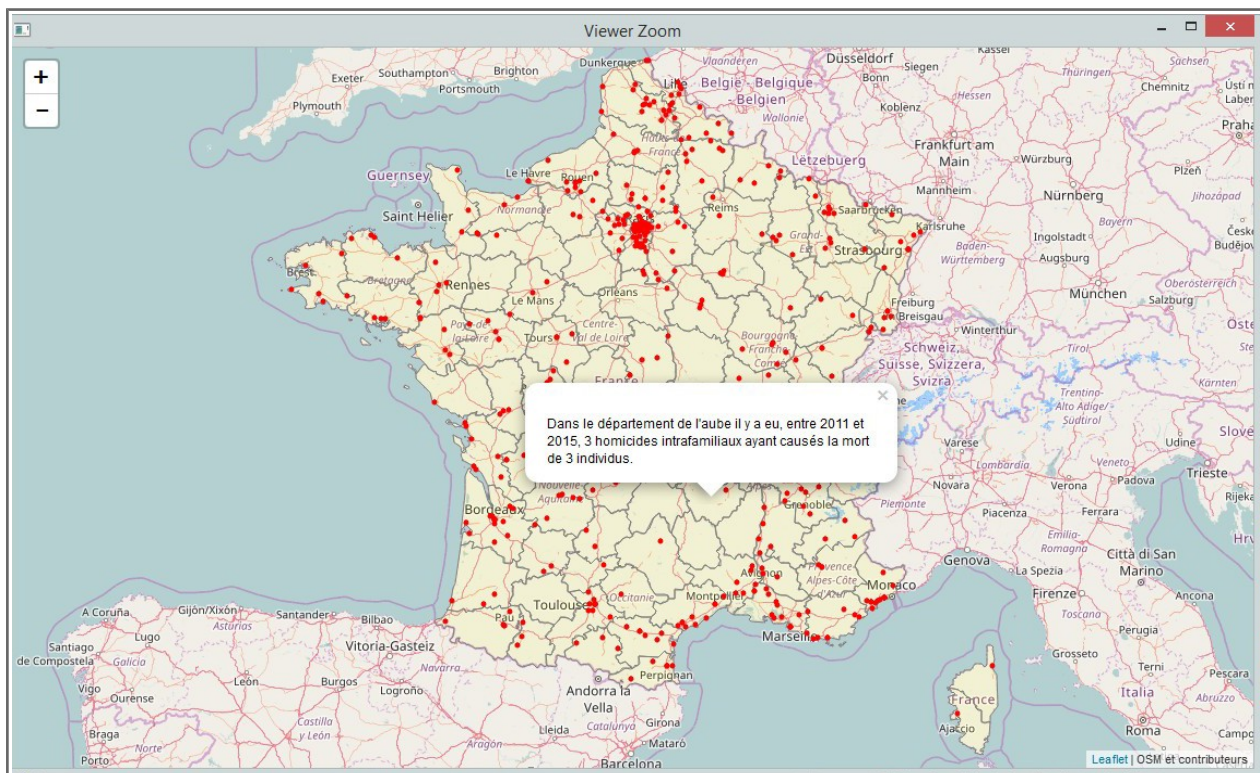
Pour mieux observer l’Ile-de-France, nous avons extrait les départements des Hautes-de-Seine, de l’Essonne, de la Seine-et-Marne, de la Seine-Saint-Denis, du Val D’Oise, du Val-de-Marne et celui des Yvelines, auxquels nous avons ajouté la ville de Paris. Nous les avons affichés dans une nouvelle fenêtre pour mieux visualiser la répartition des homicides.

Sur ces deux premières cartes, les informations comprises dans notre base de données ne sont pas toutes accessibles. Nous visualisons l’année de commission des homicides ainsi que le lieu mais nous ne connaissons ni l’âge, ni le sexe de l’auteur, ni le mode opératoire, ni le nombre de victimes. Nous avons souhaité améliorer notre travail en créant une carte interactive. Le but est :

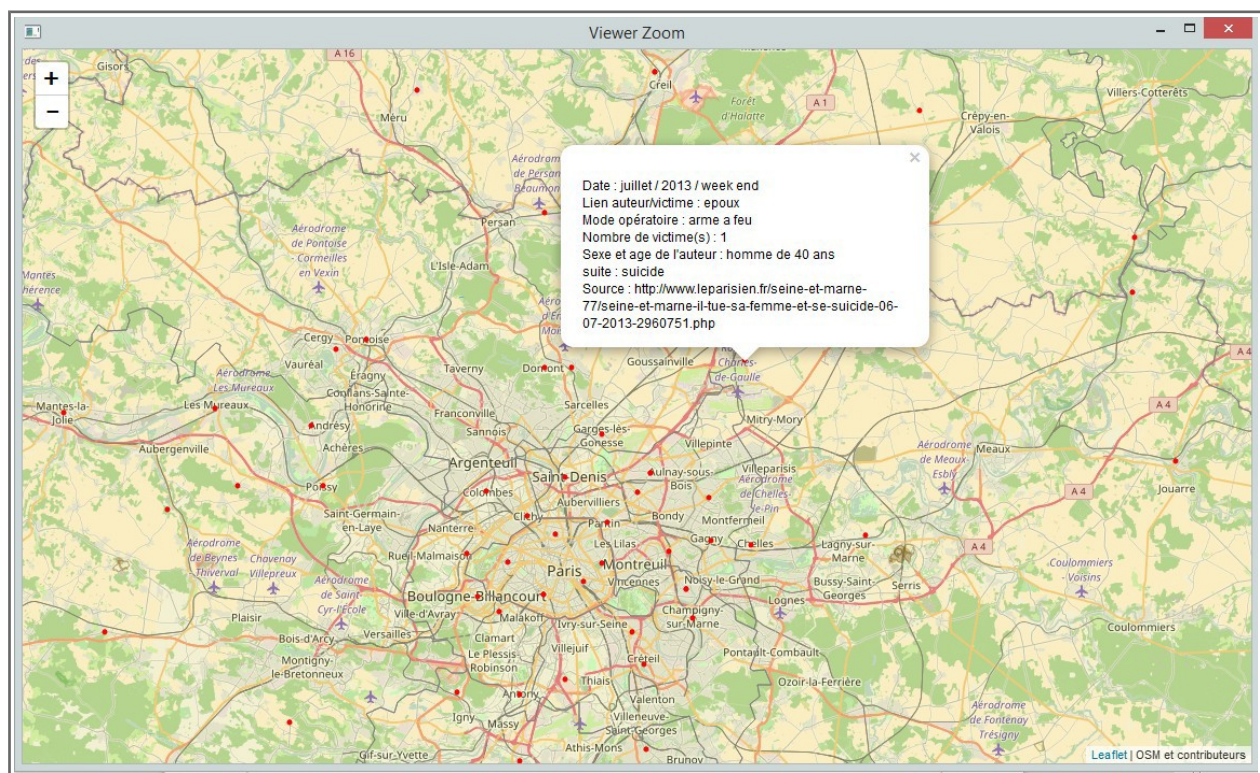
- de pouvoir zoomer sur une région ;
- d’obtenir les informations comprises dans notre fichier lorsque nous cliquons sur un point représentant un événement.

Cela est possible avec le package « leaflet », qui fait apparaître en haut à gauche les symboles « + » et « - » permettant de zoomer. Le fond de carte utilisé est celui de OpenStreetMap (OSM).

Cela est possible avec le package « leaflet », qui fait apparaître en haut à gauche les symboles « + » et « - » permettant de zoomer. Le fond de carte utilisé est celui de OpenStreetMap (OSM).



Carte interactive – Information par département



Carte interactive – Zoom sur un drame familial

```

#----- Code Afficher une carte « simple » des drames familiaux -----#

# Appel des librairies
library(cartography), library(rgdal), library(maptools), library(rgeos), library(sp),
library(ggplot2)

# Importer la carte
fr1 <- readOGR("FRA_adm1.shp", layer="FRA_adm1") #Contours

# Importer le fichier au Format CSV - Vérifier le répertoire de travail
drame2011 <- read.csv("drame2011_geocoded.csv", header = TRUE, sep = ",", dec=".",
quote="\"", fill=TRUE)

# Création d'un dataframe
data.frame(drame2011)

# ----- Affichage de la carte -----#
# Colore le fond de carte bleu mer (doit être placé au début du code)
par(bg = "#A6CAE0")

# Mise en page : création d'une marge
opar <- par(mar = c(1,1,1.2,1))

# Affichage régions bordure grise fond beige
plot(fr1, col = "#E3DEBF", border = "grey40" )

# n est égale au nombre de lignes du tableau de données
n = nrow(drame2011)

# Affichage de chaque drame en fonction de ses latitude et longitude
for ( i in ( 1 : n )) # n varie de 1 jusqu'à la dernière ligne du tableau
{ points(drame2011$longitude, drame2011$latitude, pch=17, col="#40A497", cex=1)
  i=i+1 }

# Échelle : Nous devons cliquer pour la positionner là où nous le souhaitons
cat("Cliquer là où vous voulez la barre d'échelle"); l = locator(n=1)
scalebar(d = 100, xy = c(l$x, l$y), type = "bar", below = "km", lwd = 4, divs = 2, col = "black",
cex = 0.75, lonlat = T)

# Habillage complet de la carte (fonctionne avec la librairie cartography)
layoutLayer(title = "Drames Familiaux",
  author = "S.Lefebvre / O. Robert / O. Roger", sources = "Source : Article de presse
disponibles en ligne",
  frame = TRUE, col = NA, scale = NULL, coltitle = "black",
  north = TRUE)

```

```

# ----- CARTE ZOOM ILE-DE-FRANCE -----
# isoler la région Ile-de-France dans l'objet idf
idf <- fr1[fr1$CCA_1 == 11, ]
paris <- fr3[fr3$ID_2 == 14, ]
essonne <- fr3[fr3$ID_2 == 12, ]
HSeine <- fr3[fr3$ID_2 == 13, ]
SMarne <- fr3[fr3$ID_2 == 15, ]
SSDenis <- fr3[fr3$ID_2 == 16, ]
VOise <- fr3[fr3$ID_2 == 17, ]
VMarne <- fr3[fr3$ID_2 == 18, ]
Yvelines <- fr3[fr3$ID_2 == 19, ]
plot(idf)

# Gestion des marges de la figure
par(mar = c(0,0,1.2,3.5))

# Ajouter le zoom sur Ile-de-France
plot(idf, border = "grey40")
plot(paris, border = "grey60", add = T)
plot(essonne, border = "grey60", add = T)
plot(HSeine, border = "grey60", add = T)
plot(SMarne, border = "grey60", add = T)
plot(SSDenis, border = "grey60", add = T)
plot(VOise, border = "grey60", add = T)
plot(VMarne, border = "grey60", add = T)
plot(Yvelines, border = "grey60", add = T)

```



```

# ----- CARTE INTERACTIVE -----#

# Charger la librairie cartography pour utiliser les fonctions associées
Library (cartography)

# Charger la librairie leaflet pour rendre la carte interactive
Library (leaflet)

# Importer le fichier au Format CSV - Vérifier le répertoire de travail
drametotal <- read.csv("drame_total_geocoded.csv", header = TRUE, sep = ",", dec=".", quote="\"",
fill=TRUE)

# Création d'un dataframe
data.frame(drametotal)

# Associer le fichier .shp à un objet
fr1<-readOGR("FRA_adm1.shp",layer="FRA_adm1")
france <- spTransform(fr1, "+init=epsg:4326")

# Créer du texte dans les pop-up (fenêtre interactive) lorsque l'on clique sur le point rouge
drametotal$label <- paste0 (
  "<br>", "Date : ", drametotal$mois, " / ", drametotal$annee, " / ",
  drametotal$moment_semaine,
  "<br>", "Lien auteur/victime : ", drametotal$lien_parente,
  "<br>", "Mode opératoire : ", drametotal$mode_operatoire,
  "<br>", "Nombre de victime(s) : ", drametotal$nombre_victime,
  "<br>", "Sexe et age de l'auteur : ", drametotal$sexe_auteur," de ",
  drametotal$age_auteur, " ans",
  "<br>", "suite : ", drametotal$suite,
  "<br>", "Source : ", drametotal$source
)

# Initialiser la carte
m <- leaflet()

# Ajouter un fond de carte
m <- addTiles(map = m, attribution = "OSM et contributeurs")

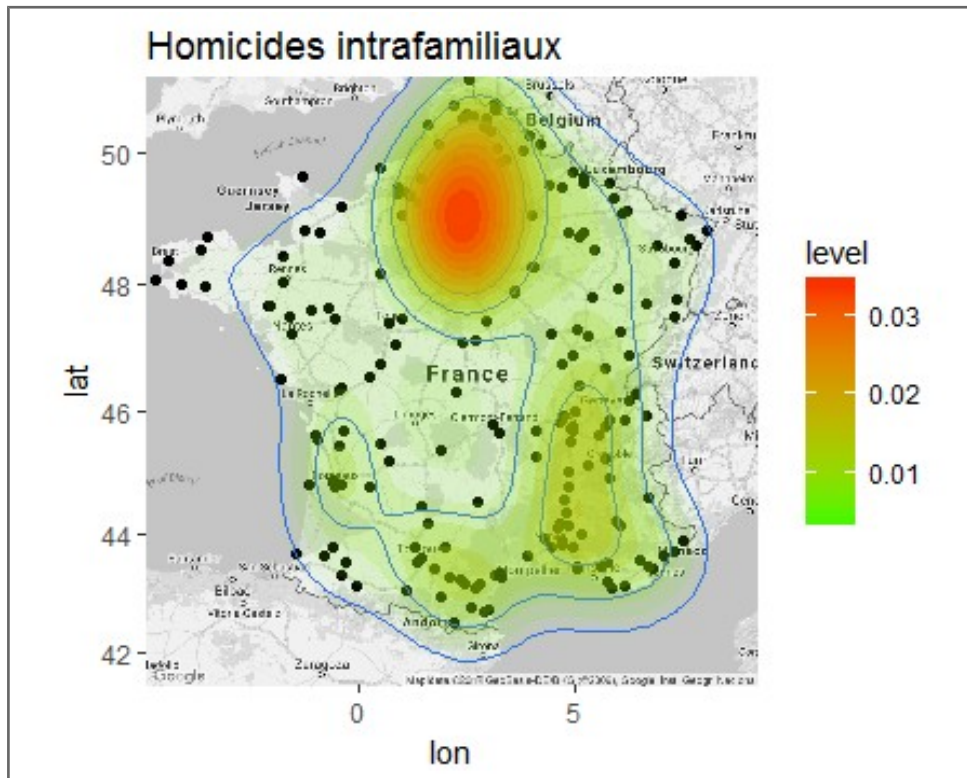
# Ajouter une couche de gones clickable affichage les départements
m <- addPolygons (map = m, data = france, Opacity = 3, color = "grey", stroke = TRUE,
  weight = 1, popup = france$label, fill = T, fillColor = "yellow", fillOpacity = 0.1)

# Affichage des drames familiaux : points rouges cliquables !
m <- addCircles(map = m, lng = drametotal$longitude, lat = drametotal$latitude,
  opacity = 3, color = "red", stroke = TRUE, weight = 3, popup = drametotal$label,
  fill = T, fillColor = "green", fillOpacity = 0.2)

m

```

A partir de nos données, nous avons réalisé une carte de densité grâce une fonction de package ggmap ainsi qu'un fond de carte issu de google. Le but étant de matérialiser des hotspots d'homicides intrafamiliaux.



```
#----- HotSpot avec GGMAP -----

# Création de la carte CENTREE de France centrée
france <- get_map ( location = c(lon = 2.12639, lat = 46.48361), zoom = 6,
  maptype = ("roadmap"), source = c("google"), crop = TRUE,
  color = ("bw"), messaging = FALSE)

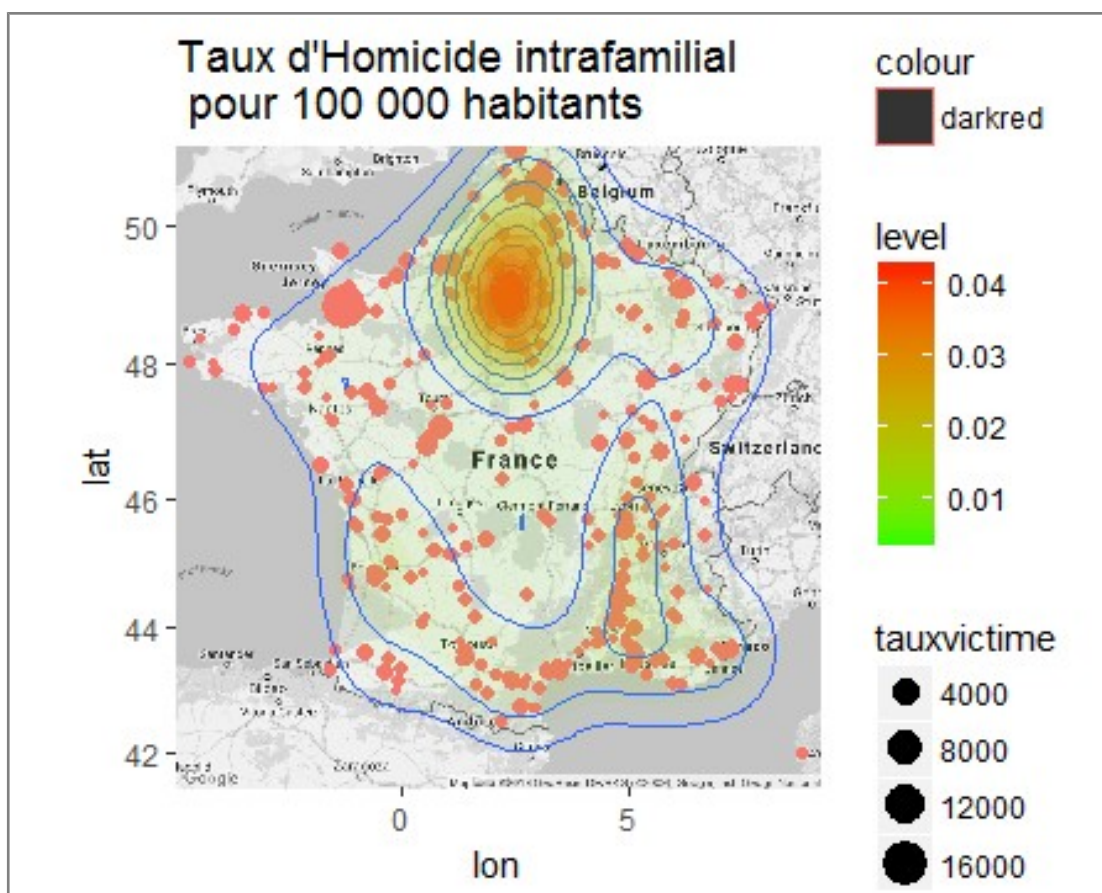
# Construction carte : Carte de France (fond) + points
dst <- ggmap (france) +
  geom_point (data= drametotal, aes(x=longitude, y=latitude))
ggtitle ("Homicides intrafamiliaux")

# Densité et affichage de la carte
dst + geom_density2d(data = drametotal,
  aes(x = longitude, y = latitude), size = 0.3) +
  stat_density2d(data = drametotal,
    aes(x = longitude, y = latitude, fill = ..level.., alpha = ..level..),
    size = 0.01, bins = 16, geom = "polygon") +
  scale_fill_gradient(low = "green", high = "red") +
  scale_alpha(range = c(0, 0.8), guide = FALSE)
```

Le résultat obtenu n'est pas explicite. M. Duchesne Sébastien, nous a conseillé de rapporter les faits à la population, afin d'obtenir un taux d'homicide. Avec M. Duchesne, nous avons également abordé la comptabilisation des suicides : dans notre étude, nous n'avons pas inclus les auteurs qui se sont suicidés dans le nombre de victimes, or, d'un point de vue criminologique, nous aurions dû étudier plus sérieusement cette possibilité.

Dans un premier temps, nous avons calculer un taux d'homicide à l'échelle de la commune. Pour se faire, nous avons appliquer la formule suivante¹⁴ :

$$\text{taux d' homicide} = \frac{\text{Nombre d' homicides par commune}}{\text{Nombre d' habitants par commune}} * 100\,000$$



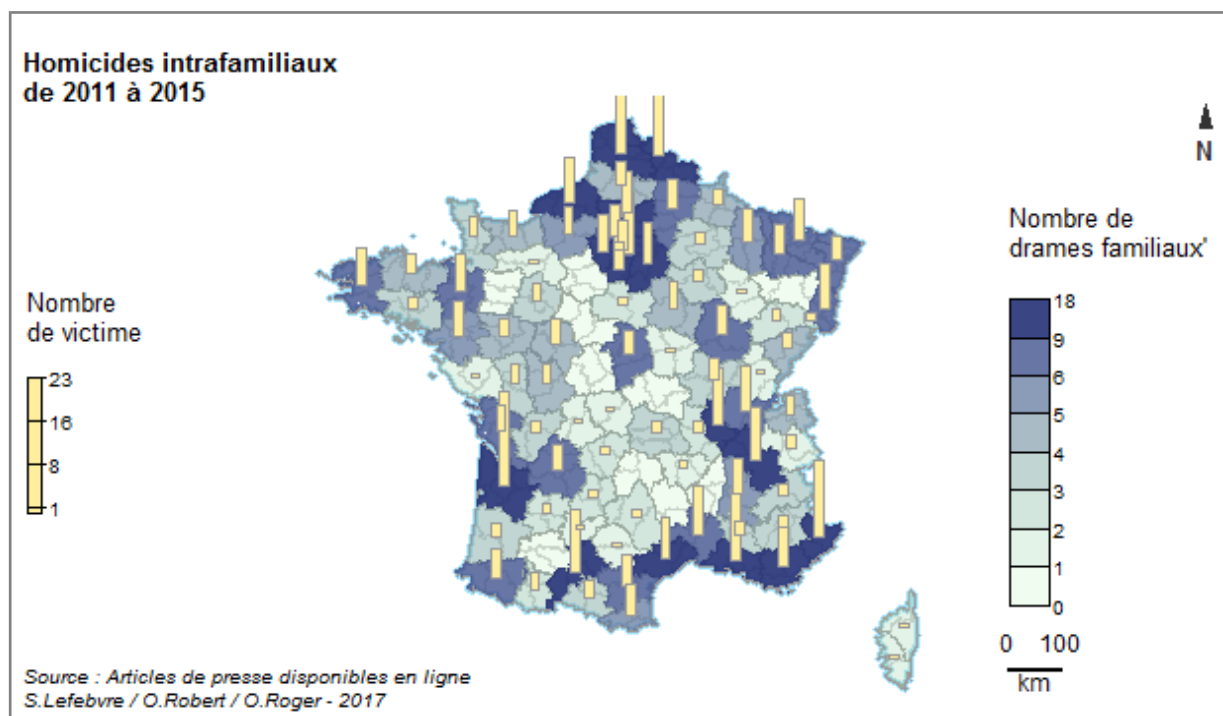
Nos connaissances en calcul de densité spatiale ne sont actuellement pas suffisantes pour exploiter pleinement l'outil existant et en donner une interprétation pertinente. Toutefois, il semble clair que les homicides intrafamiliaux prédominent largement en région parisienne.

Pour la suite de notre étude, nous avons rapporté nos homicides à l'échelle départementale.

14 Méthodes quantitatives pour les sciences humaines Par Réjean Huot page 7

Pour cela nous avons créé une seconde base de données contenant : le département, le nombre d'auteur, le nombre de victime, ainsi que la population par département, information trouvée sur le site de l'INSEE¹⁵. C'est ce nouveau fichier qui sera utilisé pour la suite de notre étude.

drame_par_dep_geocoded3.csv - LibreOffice Calc												
Fichier Édition Affichage Insertion Format Feuille Données Outils Fenêtre Aide												
Libération Sans 10												
A21:AMJ22 2b												
	A	B	C	D	E	F	G	H	I	J	K	L
1	num_dep	dep	acte_par_dep	vic_par_dep	moyenne_vic	latitude	longitude	result_label	result_score	result_type	result_id	result_housenur
2	1	ain	7	11	1.57142857142857	46.192435	5.210533	Rue Girod de l'Ain 01000 Bourg-en-Bresse	0.29	street	01053_XXX_502a35	
3	2	aisne	7	7	1.49563059	3.623119	Laon		0.3	municipality		2408
4	3	allier	0	0	0.46565333	3.333046	Rue d'Allier 03000 Moulins	0.65	street	03190_XXX_b6583		
5	4	alpes haute proven	3	3	1.44090783	6.234333	Rue de Provence 04000 Digne-les-Bains	0.39	street	04070_1350_eb127a		
6	5	hautes alpes	3	3	1.44562918	6.084359	Avenue des Alpes 05000 Gap	0.39	street	05061_XXX_295d2a		
7	6	alpes maritimes	16	19	1.187543	7.275846	Avenue des Arènes de Cimiez 06000 Nice	0.25	street	06088_XXX_dd5d0d		



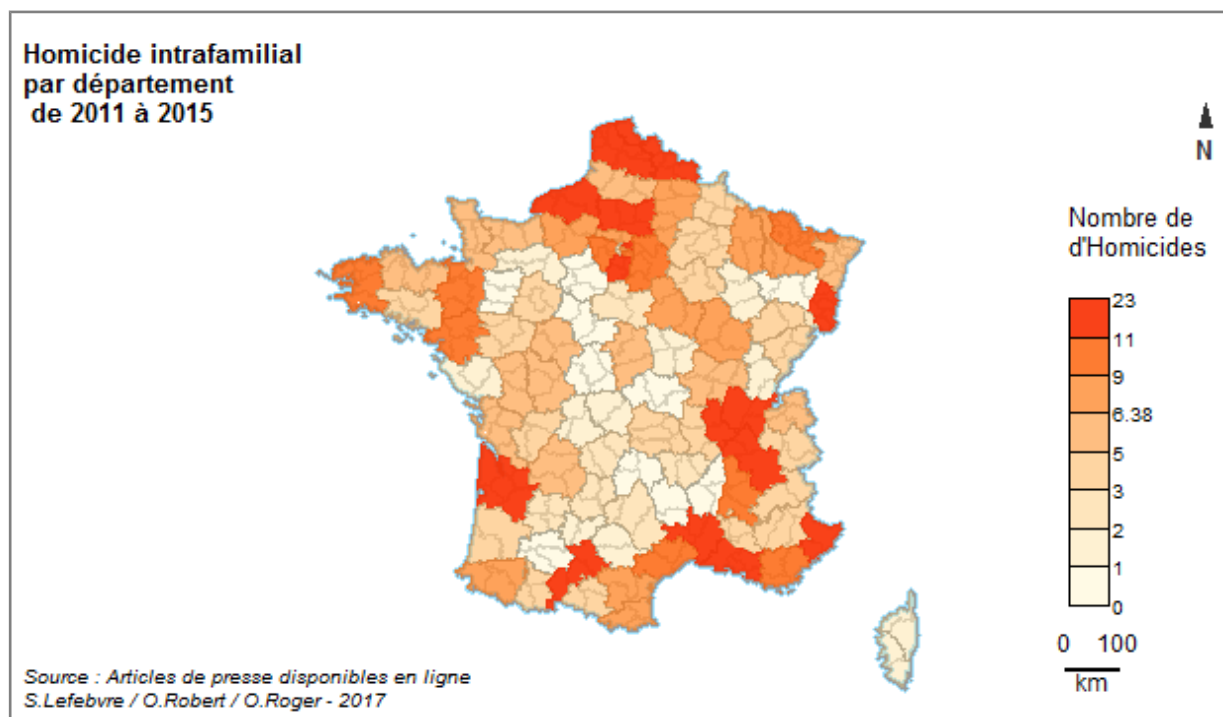
Enfin notre base de données nous a permis de construire une carte de type choroplèthe, représentant le nombre de drames familiaux par département, à laquelle nous avons ajouté le nombre de victime totale par département, sous forme de barres. Pour se faire, nous avons utilisé le package cartography.

Nous avons quelques soucis d'affichage concernant quelques départements : Ain, Aisne, Allier, Hautes-Alpes, Alpes-Maritimes, Ardèche, Ardennes, Alpes-de-haute-Provence, Corse, et l'Ariège. Même en changeant de fond de carte, le problème persiste. Nous avons compris que c'est département ont un numéro qui pose problème car Excel formate automatiquement les nombres,

¹⁵ <https://www.insee.fr/fr/statistiques/2119468?sommaire=2119504>

ainsi, par défaut, il efface le 0 devant un nombre. Les départements apparaissaient de 1 à 9 au lieu de 01 à 09. Lorsque nous utilisons la fonction « merge » pour lier deux fichiers, les départements n'étaient pas reconnus. Il a fallu formater les cellules de Excel pour corriger l'erreur, idem pour la Corse, département qu'il faut numéro 2A et 2B et non 2a et 2b.

Nous concluons que se sont les lieux où il y a le plus de drames, qu'il y a le plus de victimes, cela corrobore le fait qu'un drame intrafamilial concerne très majoritairement des homicides conjugaux et très peu d'auteur déciment sa famille entière.



Enfin, si nous comparons cette carte à celle des homicides, réalisée grâce à cartocrime (présentée en page 12), nous pouvons dire que les départements où le nombre d'homicide intrafamilial est élevé, (supérieur à 6 en 5ans) ne sont pas ceux où le nombre d'homicide est le plus élevé (exemples : les départements du Nord et du Finistère).

Il serait intéressant de comparer notre carte avec des cartes de criminalités « spécifiques » comme par exemple : celle des vols, des viols, etc...

```

# Fichier contenant les données à cartographier
dramedep <- read.csv("drame_par_dep_geocoded3.csv", header=TRUE, sep="," , dec="," , quote="\"",
fill=TRUE)
data.frame(dramedep)

# Création d'une jointure entre le fichier de données (TauxPauvreté2013) et du fond de carte
departement@data <- merge( departement@data,      #Fichier carte
                             dramedep,           #Fichier données
                             by.x="code_insee",    #Colonnes références fichier carte
                             by.y="num_dep",       #Colonnes géocodage fichier données
                             all.x=TRUE)

# Créer une palette de couleur avec la fonction carto.pal
cols <- carto.pal(pal1 = "turquoise.pal",  # palette de couleur premier dégradé
                  N1 = 9, transparency = T)  # divise en 9 couleurs rendues légèrement transparentes

# Afficher les régions françaises
opar <- par(mar = c(0.1,0.1,2.1,0.1))
plot(departement, xlim = c(-5,10), ylim = c (42,51), border = "NA")
plot(fr0, border = "#9EE1FE", lwd = 3)
plot(fr3, border = "grey80", add = T)
plot(fr2, border = "grey60", add = T)

# Affichage du nombre d'homiciede
choroLayer(spdf = departement,  # SpatialPolygonsDataFrame des départements
            df = dramedep,      # data frame taux de pauvreté
            var = "acte_par_dep", # variables utilisées
            breaks = NULL, method = "quantile", # Divise les valeurs selon la méthode quantile
            col = cols,         # couleurs
            border = NA,        # couleur de la bordure des départements
            lwd = 0.5,          # largeur dela bordure
            legend.pos = "right", # position de la légende
            legend.title.txt = "Nombre de\ndrames familiaux\n", # titre de la légende
            legend.values.rnd = 2, # number of decimal in the legend values
            legend.title.cex = 0.8,
            add = TRUE)         # ajouter la couche sur la carte existante

# Barres :nombres de victimes
propSymbolsLayer(spdf = departement, df = dramedep,
                 var = "vic_par_dep", symbols = "bar", inches = 0.5, col = "#FEEE9E", legend.pos = "left",
                 legend.title.cex = 0.8, legend.title.txt = "Nombre \nde victime\n », border = "grey60")

# Echelle
cat("Cliquer là où vous voulez la barre d'échelle");l = locator(n=1)
scalebar(d = 100, xy = c(l$x,l$y), type = "bar", below = "km", lwd = 4, divs = 2, col = "black", cex =
0.75,lonlat = T)

# Habillage complet de la carte (fonctionne avec la librairy cartography)
layoutLayer(title = "Homicides intrafamiliaux \n de 2011 à 2015",
            author = "S.Lefebvre / O.Robert / O.Roger - 2017",
            sources = "Source : Articles de presse disponibles en ligne",
            frame = TRUE, col = NA scale = NULL, coltitle = "black", north = TRUE)

```

2) Hypothèse 1 : Corrélation spatiale nombre d'armes / nombre d'homicides

Nous avons constitué une base de données à partir d'articles de journaux et avons constaté que les homicides sont principalement commis par arme blanche, dont la détention n'est pas réglementée, et par arme à feu, dont la détention est réglementée. (Actuellement, nous avons les homicides au niveau local (commune) que nous comptons ramener à l'échelle départementale.)

Nos recherches ne nous ont pas permis de savoir combien d'armes à feu sont présentes officiellement sur le territoire, et encore moins comment elles sont réparties. Par contre, nous avons découvert l'existence du fichier AGRIPPA, recensant les possesseurs d'armes à feu. Ce qui nous intéresse ce n'est pas l'identité du possesseur mais le type d'arme et surtout le lieu d'enregistrement de l'arme.

La statistique criminologique étant très proche de la statistique épidémiologique nous pensions rapprocher la propagation d'un fait criminel à celle d'une maladie. Nous avons lu un article expliquant que deux chercheurs de l'université de Californie ont tenté de modéliser le nombre d'homicides par arme à feu par une approche épidémiologique¹⁶. Le nombre d'armes pouvant être assimiler à la présence d'un virus sur un territoire et l'homicide à l'apparition de la maladie, le but étant de savoir si se sont le territoires les plus "infectés", donc possédant le plus d'armes, qui présentent au final le plus fort taux d'homicides, où si il n'y a aucune corrélation entre ces deux paramètres.

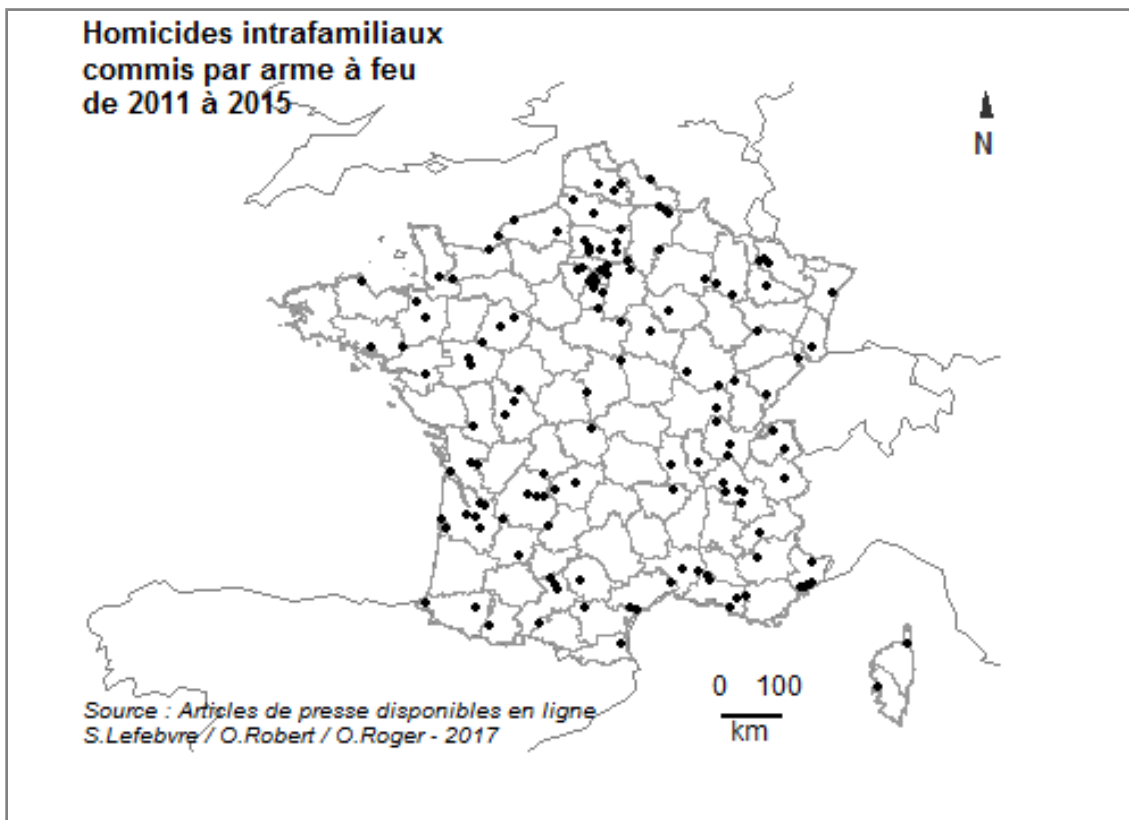
Enfin, notre base de données s'étend de 2011 à 2015, le fichier AGRIPPA n'a donc pas besoin d'avoir été actualisé après 2015 pour que nous puissions l'exploiter.

Ces données étant sensibles, nous n'avons trouvé aucune donnée en accès libre, nous avons sollicité l'aide de M. Bauer.

Ce dernier nous a répondu que cette analyse semblait difficile pour plusieurs raisons :

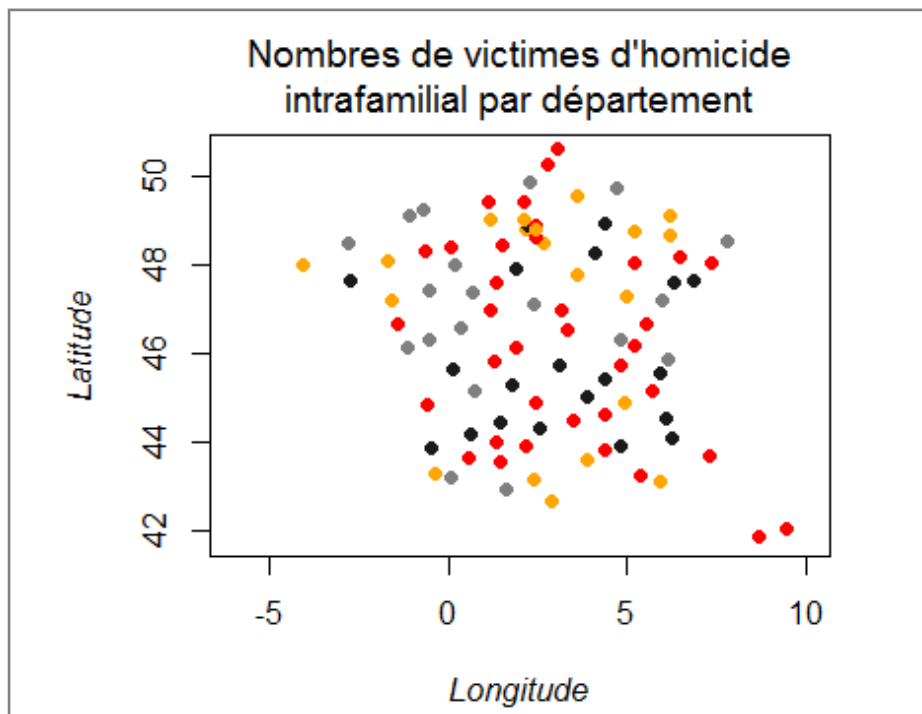
- les armes en circulation ne sont pas toutes officielles,
- la fiabilité du fichier est faible, en grande partie à cause des armes officieuses,
- les armes ne sont pas forcément utilisées sur le territoire où elles sont enregistrées.

16 <https://scienctonnante.wordpress.com/2013/09/23/pour-enrayer-la-violence-faut-il-plus-ou-moins-darmes/>



3) Normalisation du nombre de victimes par département

A partir de notre base de données contenant les homicides par département, nous avons réalisé la carte ci-dessus.



Dans un premier temps, nous avons normalisé le nombre de victimes grâce à la formule :

$$a[i] = \frac{a[i] - a_{max}}{a_{max} - a_{min}} + 1$$

avec a_{max} : le nombre maximal de victime par département ;

a_{min} : le nombre minimal de victime par département ;

nous avons ajouter 1 pour que les valeurs obtenues se situent entre 0 et 1 et non -1 et 0.

Puis, nous avons utilisé la fonction « quantile » de R afin de partager équitablement nos valeurs en 5 groupes, auxquels nous avons associé une couleur spécifique.

La « carte » graphique obtenue semble cohérente à l'exception des points situés en Corse. Il n'y a eu que 1 homicide, la couleur ne devrait pas être rouge mais noire.

```

#----- NORMALISATION DES DONNEES -----
# Nettoyage des données (inutile dans notre cas car toutes les cases comprennent une donnée exploitable)
dramedep[dramedep$vic_par_dep=="", ] <- NA
dramedep <- dramedep[complete.cases(dramedep$vic_par_dep),]
nbvic <- dramedep$vic_par_dep

# Trouver valeurs max et min : renvoie l'indice donc la ligne des valeurs extrêmes
nbvicmax <- which.max(nbvic)
nbvicmin <- which.min(nbvic)
nbvicmax
nbvicmin

# Normalisation :
n = nrow(dramedep) ; i = 0 ; nbvicnormal = 0 ;
for (i in 1:n)
{ a = dramedep$vic_par_dep[i]
  nbvicnormal[i] = ((a - dramedep$vic_par_dep[nbvicmax]) / (dramedep$vic_par_dep[nbvicmax] -
dramedep$vic_par_dep[nbvicmin]))+1 )
  i = +i}
nbvicnormal # vecteur contenant le nombre de victime normalisé
plot(nbvicnormal)

couleurs=rep(nbvicnormal)

# Division de ma série de données (du vecteur) par quantile
QuantNbVicNormal <- quantile(nbvicnormal, c(0.2,0.4,0.6,0.8))

# Création du dégradé de couleurs (1 couleur par quantile)
degrad=c("grey100","grey10","grey50","orange","red")

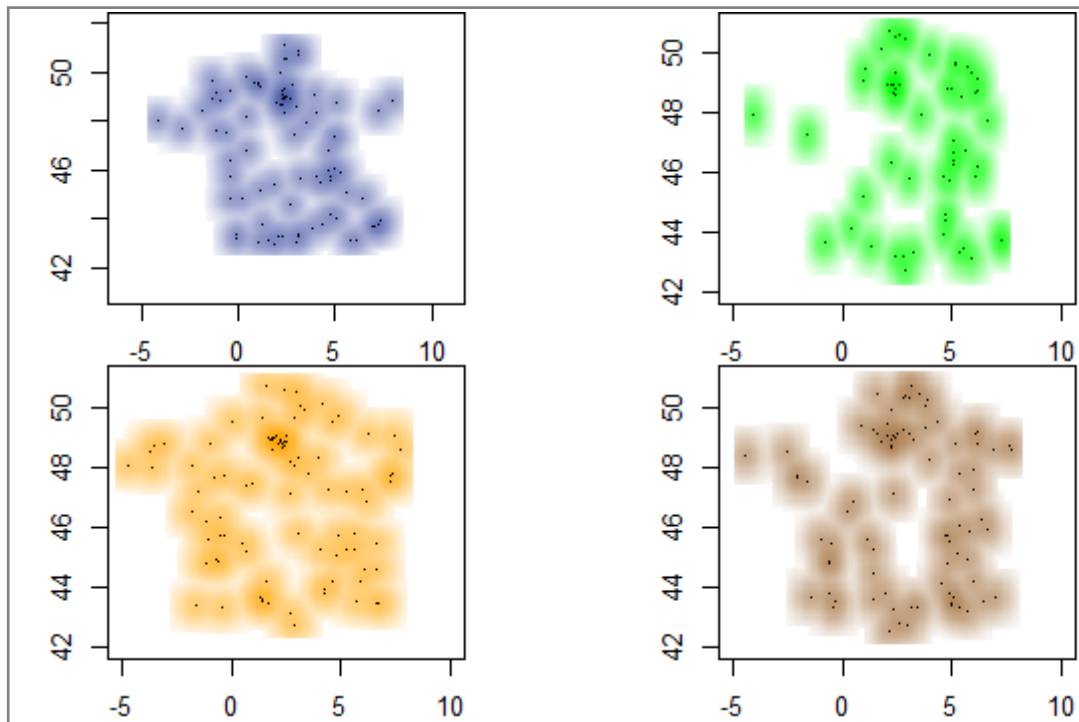
# Attribution des couleurs en fonction de la valeur des données normalisées
n = nrow(dramedep) ; i = 0 ; j = 0;
for (i in (1:n))
{ couleurs[i]="red"
  for (j in (1:4))
  {if (nbvicnormal[i]>QuantNbVicNormal[j])
    couleurs[i]=degrad[j+1]} }

# Ranger les données dans une matrice par concaténation
NbVicNorCoul <- cbind (dramedep$CODGEO, dramedep$latitude ,dramedep$longitude, nbvicnormal,
couleurs)

opar <- par(mar = c(4.2,5.2,3.2,10.2))
plot (x=dramedep$longitude, y=dramedep$latitude,
  col = couleurs, ylim = c(41.8,50.6), xlim = c(-6,10),
  font.main = 1 ,
  main = "Nombres de victimes d'homicide\nintrafamilial par département",
  sub = "Auteurs : S.L / O.R / O.R - 2017",
  font.lab = 3, pch= 16,
  xlab = "Longitude", ylab = "Latitude" )

```

4) Hypothèse 2 : Influence des saisons



Sur ces cartes sont représentées les drames familiaux par saison :

- en haut à gauche, l'hiver (janvier, février, mars),
- en haut à droite, le printemps (avril, mai, juin),
- en bas à gauche, l'été (juillet, août, septembre)
- en bas à droite, l'automne (octobre, novembre, décembre)

Elles ont été réalisées avec la fonction graphique `smoothScatter`, qui permet de calculer des densités. On constate clairement que la saison la moins meurtrière est le printemps. En été, le nombre d'homicide est élevé et surtout ils ont lieu sur l'ensemble du territoire.

```
#----- HOTSPOT -----
library(MASS)

# Diviser la fenêtre
layout(matrix(1:4,2,2))

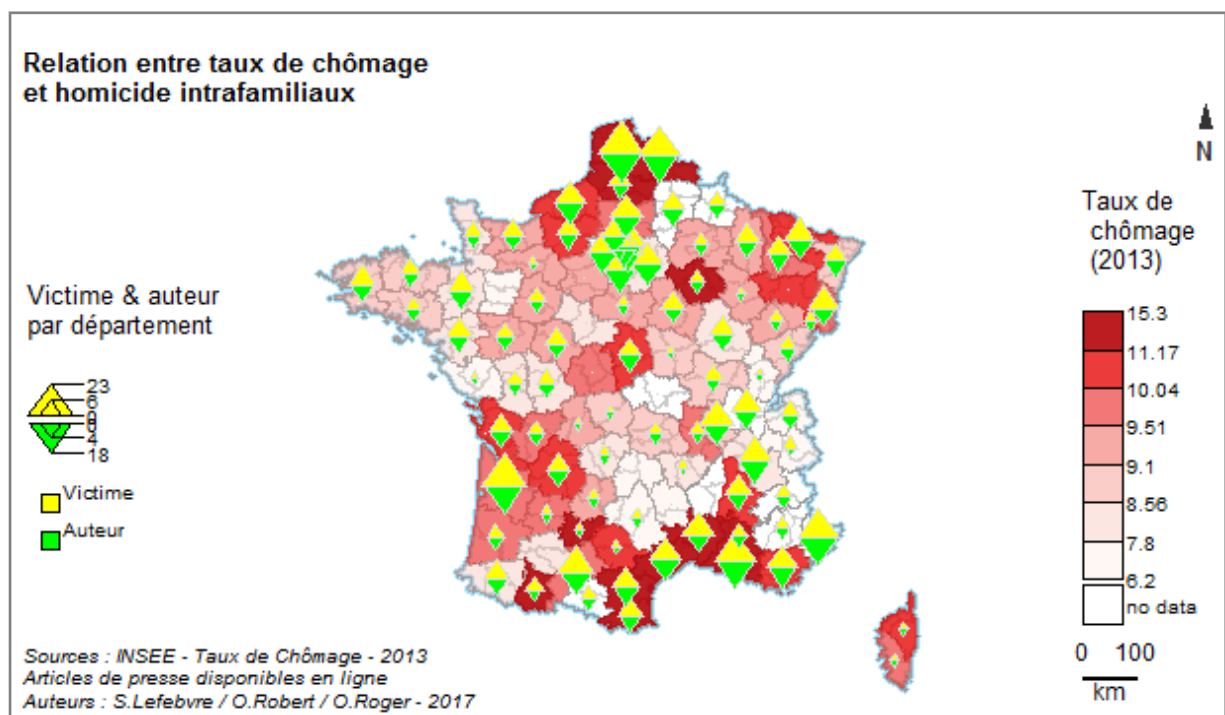
# Répartir les données par saison !
# Hiver ( la même opération sera réalisée pour les autres saisons mais ne seront pas présentés pour alléger le devoir
n = nrow(drametotal) ; drahiver = 0 ; drahiverlatitude = 0 ; drahiverlongitude = 0
for (i in 1:n)
{ drahiver = drametotal$mois[i]
  if (drahiver == "janvier" || drahiver == "fevrier" || drahiver == "mars")
  { drahiverlatitude[i] <- drametotal$latitude[i]
    drahiverlongitude[i] <- drametotal$longitude[i] }
  i = +i }

x <- drahiverlongitude
y <- drahiverlatitude
opar <- par(mar = c(2,3,0,5))
smoothScatter(x,y, xlab = "Longitude", ylab = "Latitude", xlim = c(-6, 11), ylim = c(41,52),
  colramp = colorRampPalette(c("white", "#4E5EAC")) )
```

5) Hypothèse 3 : Influence du taux de chômage «environnemental »

Lorsque nous nous sommes lancé dans la constitution de notre base de données, nous souhaitions inclure la situation professionnelle de l'auteur. Hélas, cette information est très peu présente dans les articles de journaux, elle n'est précisée que lorsqu'elle est « remarquable » comme par exemple un architecte connu pour ses travaux, un policier ayant utilisé son arme de service ou un boucher ayant utilisé un couteau spécifique à sa profession. Lorsque l'auteur a plus de 75 ans, il est légitime de considérer qu'il est retraité mais ce n'est qu'une hypothèse, un a priori dont nous n'avons pas tenu compte.

Nous avons tenter de pallier ce manque d'information par le taux de chômage à l'échelle du département. Ces données sont disponibles sur le site de l'INSEE. Nos données allant de 2011 à 2015, nous avons choisi d'utiliser les données de 2013 et avons réalisé la carte suivante :

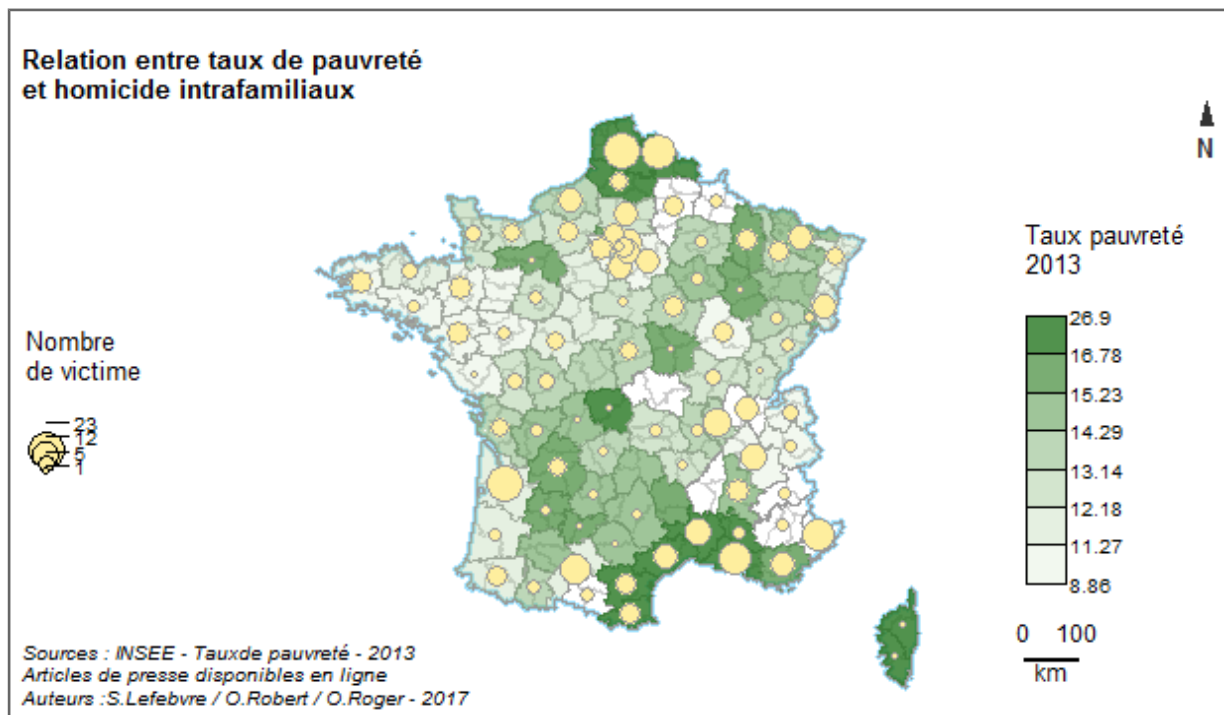


Nous constatons que le nombre de victimes est plus élevés dans les départements où le taux de chômage est supérieur à 9,1 %.

6) Hypothèse 4 : Taux de pauvreté

« Le désavantage social et économique est étroitement relié à la criminalité , en particulier aux infractions les plus graves, dont les voies de fait, le vol qualifié et les homicides »¹⁷

Après avoir étudié le taux de chômage, nous nous sommes intéressé au taux de pauvreté. Là encore les données ont été trouvées sur le site de l'INSEE. Nous avons décidé, comme pour le taux de chômage, de prendre en compte les chiffres publiés pour l'année 2013.



Notre analyse reste sensiblement la même que pour la carte précédente, le nombre d'homicide intrafamilial est le plus élevé dans les départements où le taux de pauvreté est supérieur à 13,14 %.

17 Examen du lien entre la criminalité et la situation socio-économique à Ottawa et à Saskatoon : Analyse géographique à petite échelle Peter Kitchen, Ph.D. - Mai 2006

CONCLUSION ET LIMITES

Au terme de cette étude, nous pouvons conclure que l'auteur type de l'homicide intrafamilial est un homme, âgé de 47ans, vivant dans sur un territoire où le taux de chômage est élevé (supérieur à 9,1%) et où le taux de pauvreté est lui aussi élevé (supérieur à 13,14%). Pour passer à l'acte il utilisera une arme à feu ou une arme blanche. Son crime se déroulera à plus de risque de se produire l'été, en semaine, à son domicile, contre sa femme à la suite d'une dispute. Si cet auteur ne se suicide pas après avoir commis l'homicide, il a une forte probabilité à être interpellé où immédiatement où après une courte cavale (moins de 48h).

Nous avons réussi a dresser un portrait de l'auteur ainsi que les conditions de son crime, mais les limites de cette étude sont de plusieurs natures :

- nos données ne sont pas exhaustives et dans 5 départements nous n'avons trouvé aucun homicide intrafamilial, il serait intéressant de développer l'étude sur ces territoires afin de savoir si nous sommes passé à côté de d'homicide. Si ce n'est pas le cas, il serait tout aussi intéressant de les étudier pour comprendre leur spécificité ;
- nos connaissances sont « limitées » : pour abordé le sujet, nous n'avons pas fait d'état de l'art, le logiciel R est loin d'avoir été pleinement utilisé, enfin notre analyse cartographique est perfectible ;
- enfin nous avons manqué de temps, le sujet choisi est ambitieux et mériterait un plus grand développement.

Annexe 1 : Indexation des infractions

FC2 : DÉTERMINER L'INDEX CORRESPONDANT À CHAQUE INFRACTION

A chaque infraction du champ de la statistique 4001 correspond un des 107 index. On le détermine en se référant :

- D'abord, au répertoire alphabétique des infractions qui fait correspondre un index au libellé de chaque infraction (cf 2^{ème} partie).
- Ensuite, au répertoire analytique des index qui précise le contenu et les limites de chacun d'eux (cf. 3^{ème} partie)

