

# FINAL

KASHALA ILUNGA Caleb

01/10/2019

## Table des matières

0.1	Descriptifs des variables . . . . .	2
<b>1</b>	<b>Etude sur les cycliste</b>	<b>5</b>
1.1	Age et Sexe des cyclistes . . . . .	6
1.2	Alcoolémie . . . . .	8
1.3	Race du cycliste . . . . .	9
1.4	Blésures du cycliste. . . . .	9
1.5	Relation entre type de blessure et Position du vélo . . . . .	10
<b>2</b>	<b>Etude sur les conducteurs</b>	<b>11</b>
2.1	Sexe du conducteur . . . . .	12
2.2	Age des chauffeurs. . . . .	13
2.3	Blésures des chauffeurs . . . . .	14
2.4	Relation entre la vitesse du chauffeur, sa couleur de peau et le nombre d'accidents . . . . .	15
2.5	Type de voitures . . . . .	17
<b>3</b>	<b>Etude sur les données temporelles</b>	<b>18</b>
3.1	Evolution du Nombre d'accident par mois en fonction des années . . . . .	19
<b>4</b>	<b>Carte</b>	<b>20</b>

## 0.1 Descriptifs des variables

- FID : Identifiant;
- OBJECTID : Identifiant ;
- AmbulanceR : Prise en charge par une ambulance ;
- BikeAge\_Gr : Tranche d'age du cycliste ;
- Bike\_Age : Age du cycliste ;
- Bike\_Alc\_D : Alcoolémie du cycliste ;
- Bike\_Dir : Direction du cycliste ;
- Bike\_Injur : Blessure du cycliste ;
- Bike\_Pos : Position du cycliste ;
- Bike\_Race : La race du cycliste ;
- Bike\_Sex : Le sexe du cycliste ;
- City : La Ville;
- County : Comté ;
- CrashAlcoh : Si une des deux personnes était alcoolisé ;
- CrashDay : S'il a déjà eu un accident de travail ;
- Crash\_Date : Date de l'accident ;
- Crash\_Grp : La cause de l'accident ;
- Crash\_Hour : L'heure de l'accident ;
- Crash\_Loc : S'il y avait une intersection ou pas ;
- Crash\_Mont : ;Le mois de l'accident ;
- Crash\_Time : La date de l'accident ;
- Crash\_Type : Type de l'accident ;
- Crash\_Ty\_1 ; ;
- Crash\_Year : L'année de l'accident ;
- Crsh\_Sevri : La sévérité de la blessure ;
- Development : Le cadre de la région dans lequel l'accident s'est déroulé ;
- DrvrAge\_Gr : La tranche d'age du conducteur qui a fait l'accident ;
- Drvr\_Age : L'age du conducteur ;
- Drvr\_Alc\_D : Si le conducteur était alcoolisé ou pas ;
- Drvr\_EstSp : Vitesse du conducteur responsable de l'accident ;
- Drvr\_Injur : L'état du conducteur responsable de l'accident ;
- Drvr\_Race : La race du conducteur responsable de l'accident ;
- Drvr\_Sex : Le sexe du conducteur responsable de l'accident ;
- Drvr\_VehTy : Le type de véhicule ;
- ExcsSpdInd : Depassement de la vitesse limite autorisée;
- Hit\_Run : Si le conducteur s'est enfuit sans vérifier que le pilote va bien ;
- Light\_Cond : La luminosité ;
- Locality : Localité ;
- Num\_Lanes : Numéro de la voie ;
- Num\_Units : Numéro de l'unité ;
- Rd\_Charact ; ;
- Rd\_Class : Le type de route ;
- Rd\_Conditi : Condition de la route ;
- Rd\_Config : Le genre de la route ;
- Rd\_Defects : L'Etat de la route ;
- Rd\_Feature : Spécificité de la route ;
- Rd\_Surface : La surface de la route ;
- Region : La Region ;
- Rural\_Urba : Rurale ou urbaine ;
- Speed\_Limi : Limite de vitesse ;
- Traff\_Cntr : Controle du trafic ;
- Weather : Condition météo ;

- `Workzone_I` : Zone de travail ;
- `Location` : Localisation ;

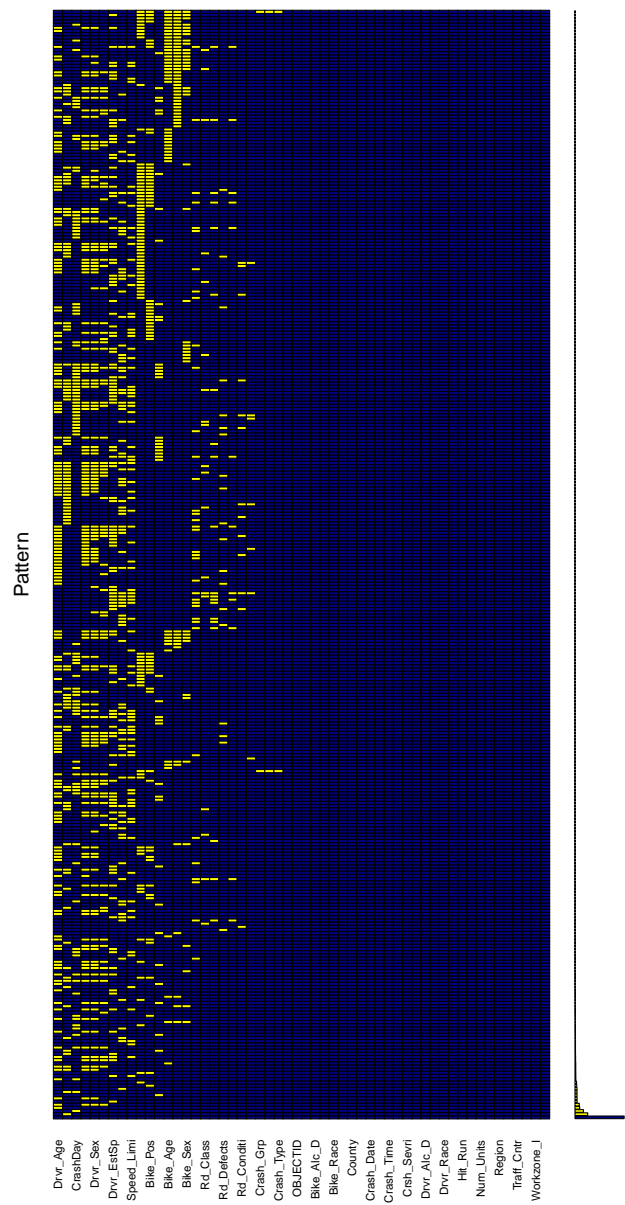
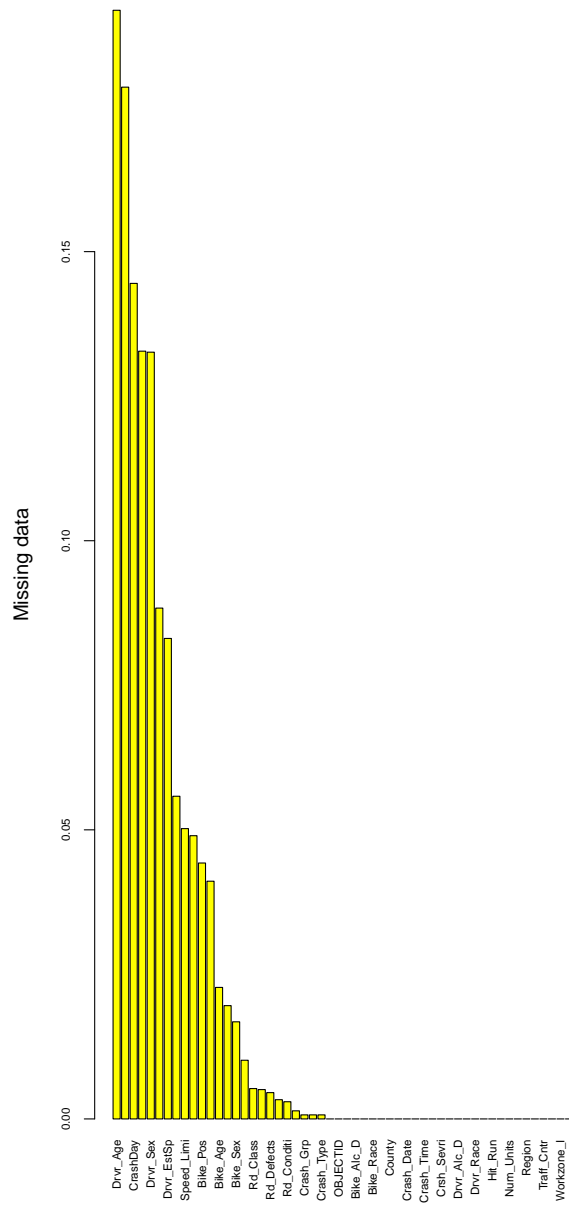
Pour notre étude analytique nous allons travailler sur une base de données qui nous donne un ensemble d'informations sur les accidents de vélo qui se sont produit aux Etats-Unis. Elle nous donne des informations sur des variables inhérentes aux conducteurs impliqués dans ces accidents mais aussi sur les divers conditions et situations physiques et environnementales qui l'entoure. Cette base de données nous renseigne sur 5716 observations contenant 54 variables.

`#Plan`

Pour mener à bien notre analyse nous allons procédés en plusieurs etapes : `* Etudes sur les cyclistes` : consacré à l'études des variables qui concerne le cycliste ; `* Etude sur les conducteurs` : consacré à l'études des variables qui concerne les conducteurs ; `* Etudes sur les données temporelles` : consacré à l'études des variables qui concerne le cycliste ; `* Etude sur les condtions de l'accident` : qui correpond aux données sur les conditions météo, l'état de la route etc.. ;

- `Etude sur les autres variables` : qui correspond aux données non classées mais qui semblent pertinentes ;

`#Données manquantes`



```
##
## Variables sorted by number of missings:
## Variable      Count
## Drvr_Age 0.1917424773
## Crash_Ty_1 0.1784464661
## CrashDay 0.1445066480
## DrvrAge_Gr 0.1327851645
## Drvr_Sex 0.1326102169
## Drvr_VehTy 0.0883484955
## Drvr_EstSp 0.0831000700
## Num_Lanes 0.0558082575
## Speed_Limi 0.0502099370
## Bike_Dir 0.0489853044
## Bike_Pos 0.0442617215
```

```

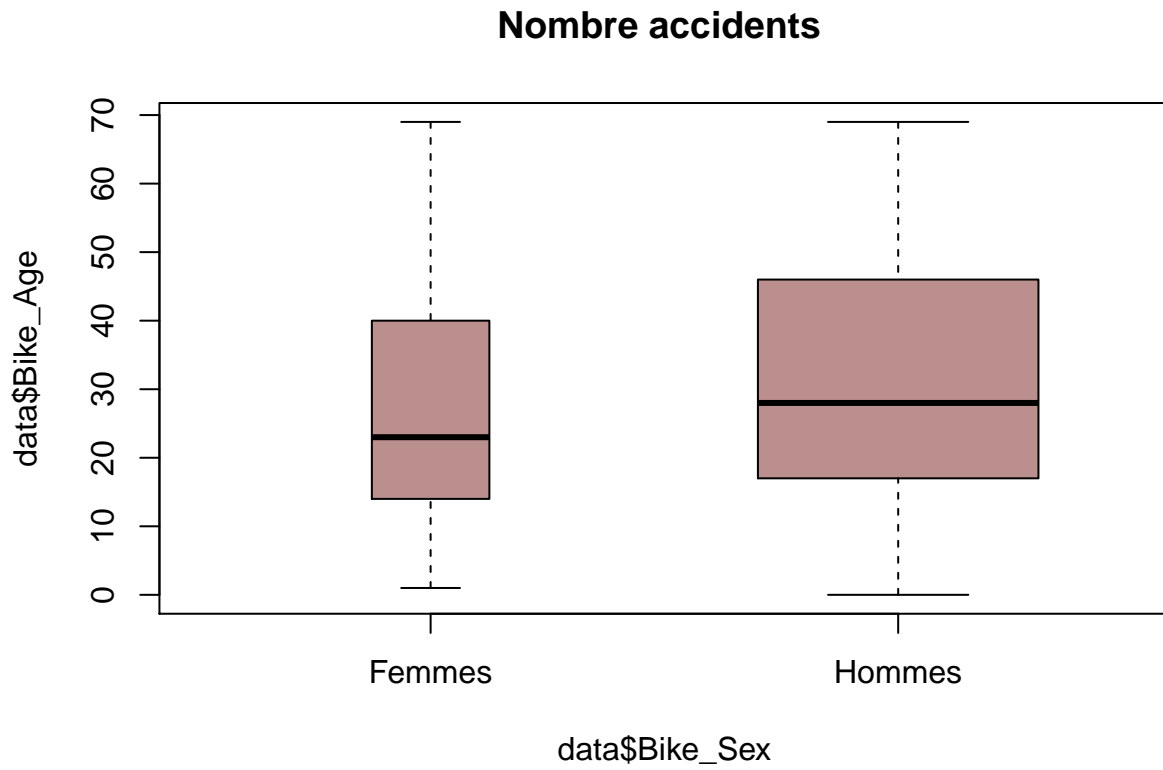
## Crash_Mont 0.0411126662
## Bike_Age 0.0227431770
## BikeAge_Gr 0.0195941218
## Bike_Sex 0.0167949615
## Rd_Config 0.0101469559
## Rd_Class 0.0052484255
## Rd_Charact 0.0050734780
## Rd_Defects 0.0045486354
## Rd_Surface 0.0033240028
## Rd_Conditi 0.0029741078
## Light_Cond 0.0013995801
## Crash_Grp 0.0006997901
## Crash_Loc 0.0006997901
## Crash_Type 0.0006997901
## FID 0.0000000000
## OBJECTID 0.0000000000
## AmbulanceR 0.0000000000
## Bike_Alc_D 0.0000000000
## Bike_Injur 0.0000000000
## Bike_Race 0.0000000000
## City 0.0000000000
## County 0.0000000000
## CrashAlcoh 0.0000000000
## Crash_Date 0.0000000000
## Crash_Hour 0.0000000000
## Crash_Time 0.0000000000
## Crash_Year 0.0000000000
## Crsh_Sevri 0.0000000000
## Developmen 0.0000000000
## Drvr_Alc_D 0.0000000000
## Drvr_Injur 0.0000000000
## Drvr_Race 0.0000000000
## ExcsSpdInd 0.0000000000
## Hit_Run 0.0000000000
## Locality 0.0000000000
## Num_Units 0.0000000000
## Rd_Feature 0.0000000000
## Region 0.0000000000
## Rural_Urba 0.0000000000
## Traff_Cntr 0.0000000000
## Weather 0.0000000000
## Workzone_I 0.0000000000
## Location 0.0000000000

```

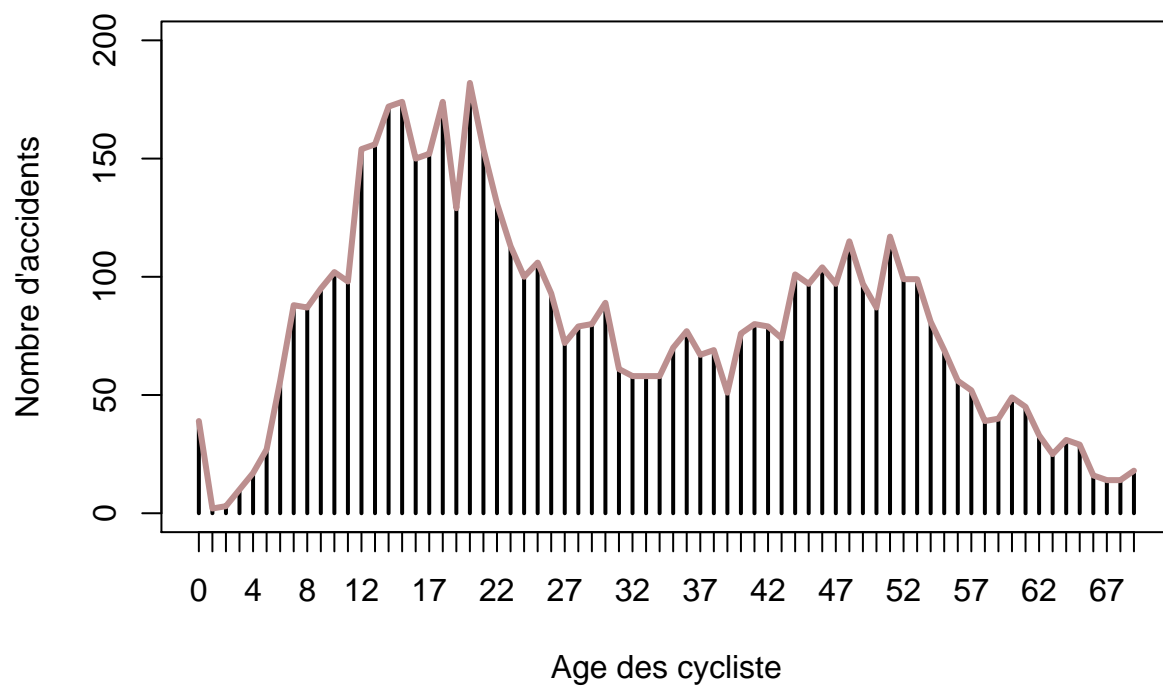
Les variables qui contiennent le plus de données manquantes sont : L'age du vhauffeur, Crash\_ty\_1, Le jour de du crash , et le sexe du chauffeur.

## 1 Etude sur les cycliste

## 1.1 Age et Sexe des cyclistes



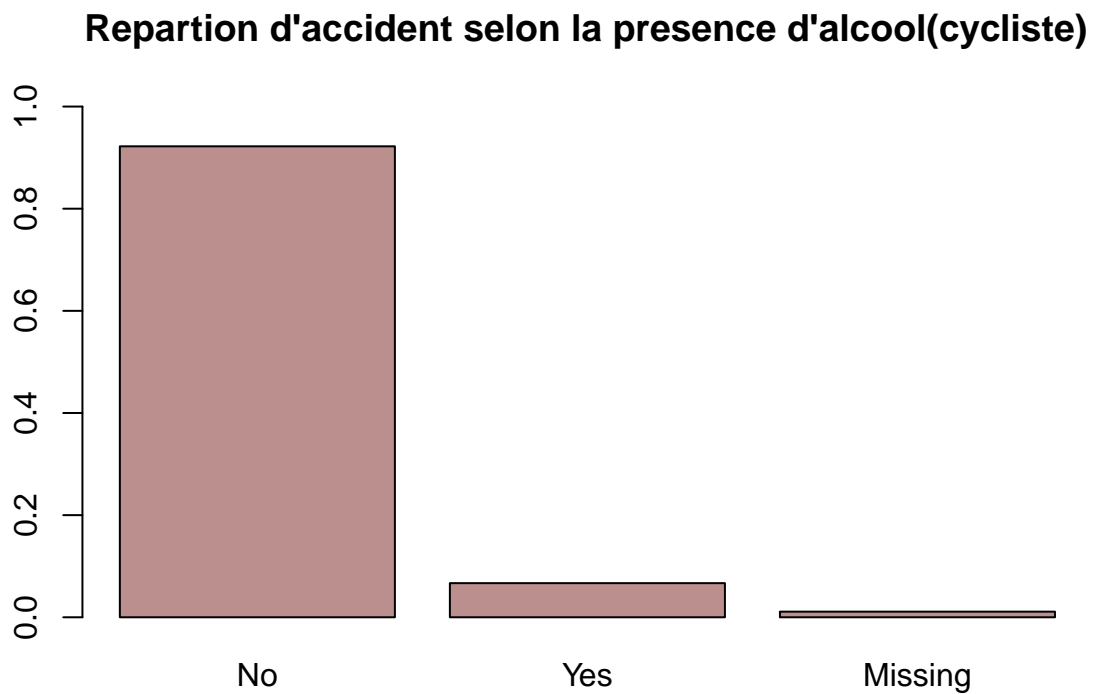
Nous remarquons sur ce graphique que les accidents de vélo concernent essentiellements les hommes et que la moyennes d'age des accidentées etait relativement la meme quelque soit le sexe.



L'âge des accidentés varie fortement, cependant on peut constater sur ce graphique que le nombre d'accidents est beaucoup plus élevé pour les individu ayant entre 13 et 21 ans ; Puis s'en suis une forte baisse du nombre d'accidents chez les 22-40 ans puis, après une légère augmentation chez les 41-53 ans, le nombre d'accident se réduit considérablement.

	nombre	Proportion
White	3111	0.5442617
Other	48	0.0083975
native American	67	0.0117215
Hispanic	297	0.0519594
Black	2006	0.3509447
Asian	56	0.0097971
Missing	131	0.0229181

## 1.2 Alcoolémie

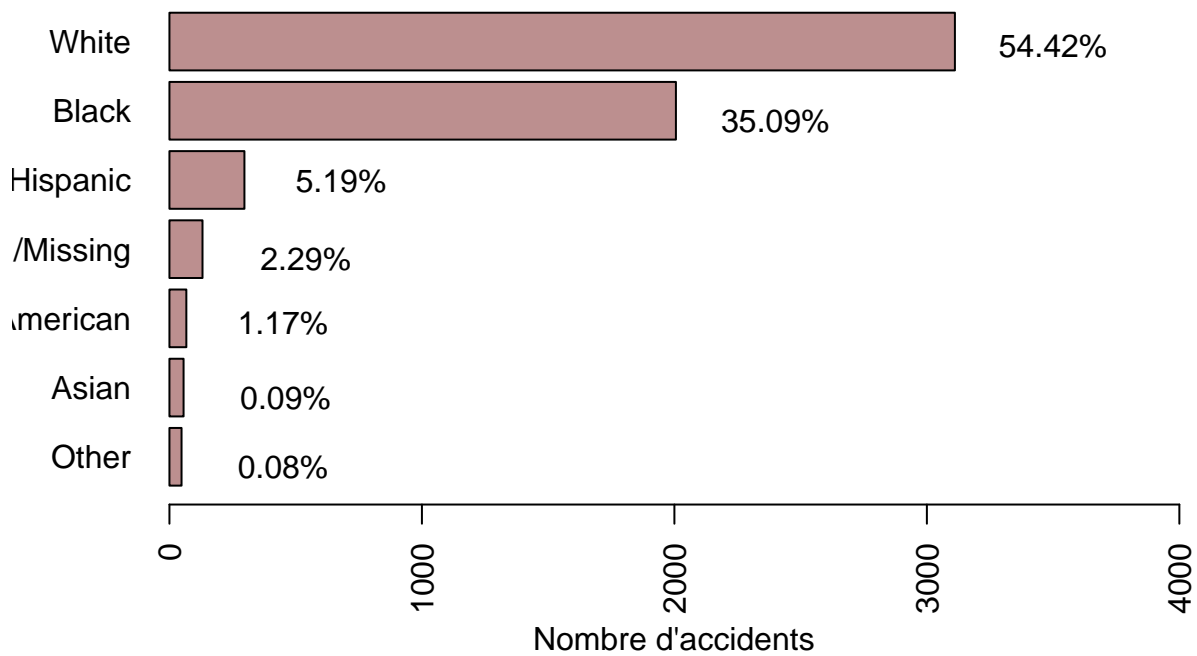


La plupart des accidents ne sont pas dû à l'alcool.



Bike_Injur	Freq
B: Evident Injury	2405
C: Possible Injury	2199
O: No Injury	526
A: Disabling Injury	291
Injury	172
K: Killed	123

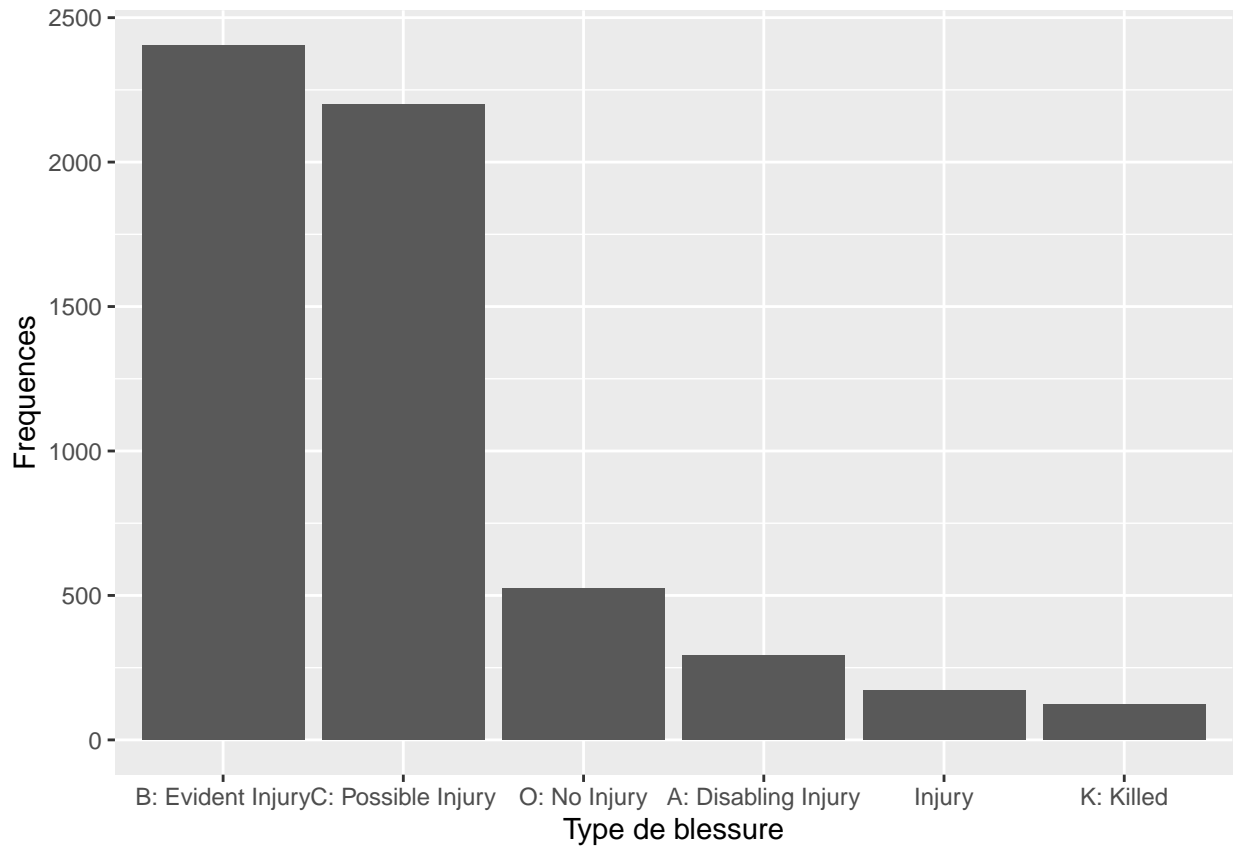
### 1.3 Race du cycliste



### 1.4 Blessures du cycliste.

```
## Warning: Use of `Injur$Bike_Injur` is discouraged. Use `Bike_Injur` instead.
## Warning: Use of `Injur$Freq` is discouraged. Use `Freq` instead.
```

Var1	Freq
Travel Lane	3746
Sidewalk / Crosswalk / Driveway Crossing	899
Non-Roadway	309
Bike Lane / Paved Shoulder	295
Driveway / Alley	146
Multi-use Path	36
Other	32



La majeure partie des accidents, bien que n'étant pas mortel, entraîne dans la plupart des cas des blessures graves. Ceci s'explique notamment parce que la plupart des collisions se font avec des automobilistes et que les cyclistes n'ont aucune protection d'aucune sorte.

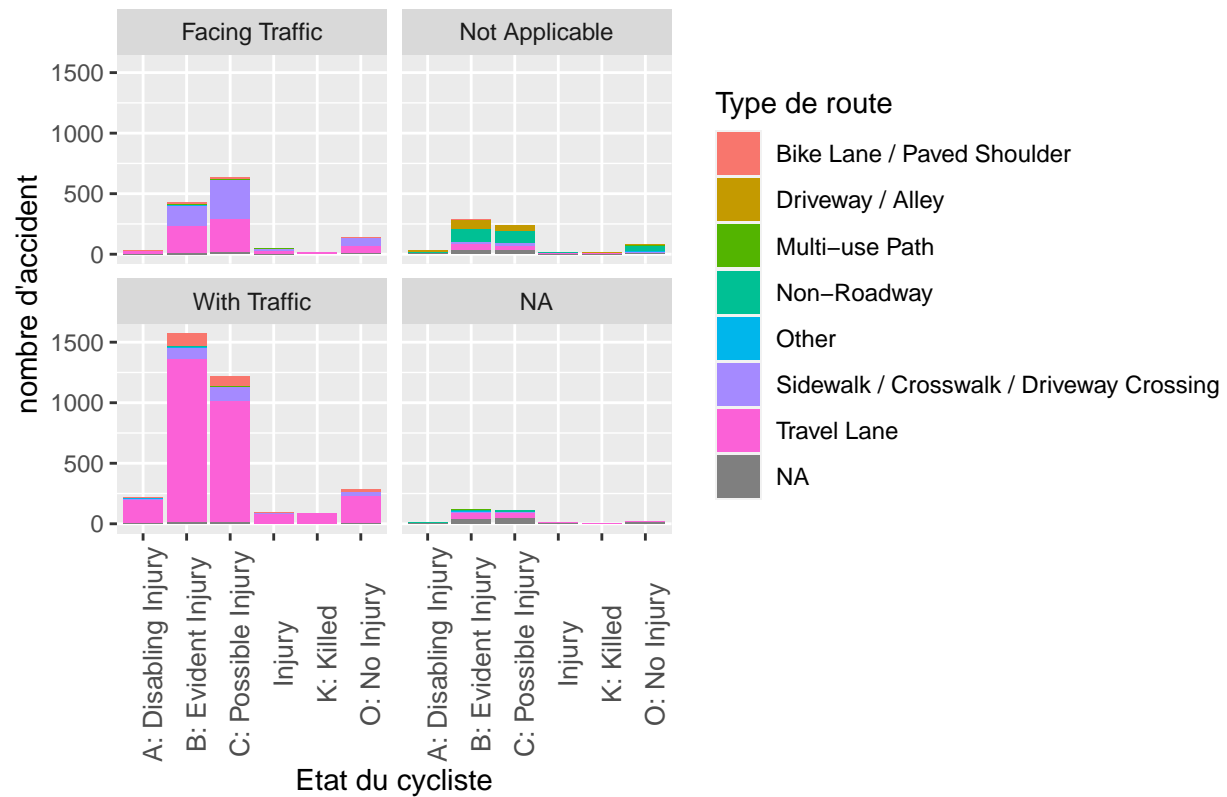
### 1.5 Relation entre type de blessure et Position du vélo

```
## Warning: Use of `bike$Bike_Injur` is discouraged. Use `Bike_Injur` instead.
```

```
## Warning: Use of `bike$Bike_Pos` is discouraged. Use `Bike_Pos` instead.
```

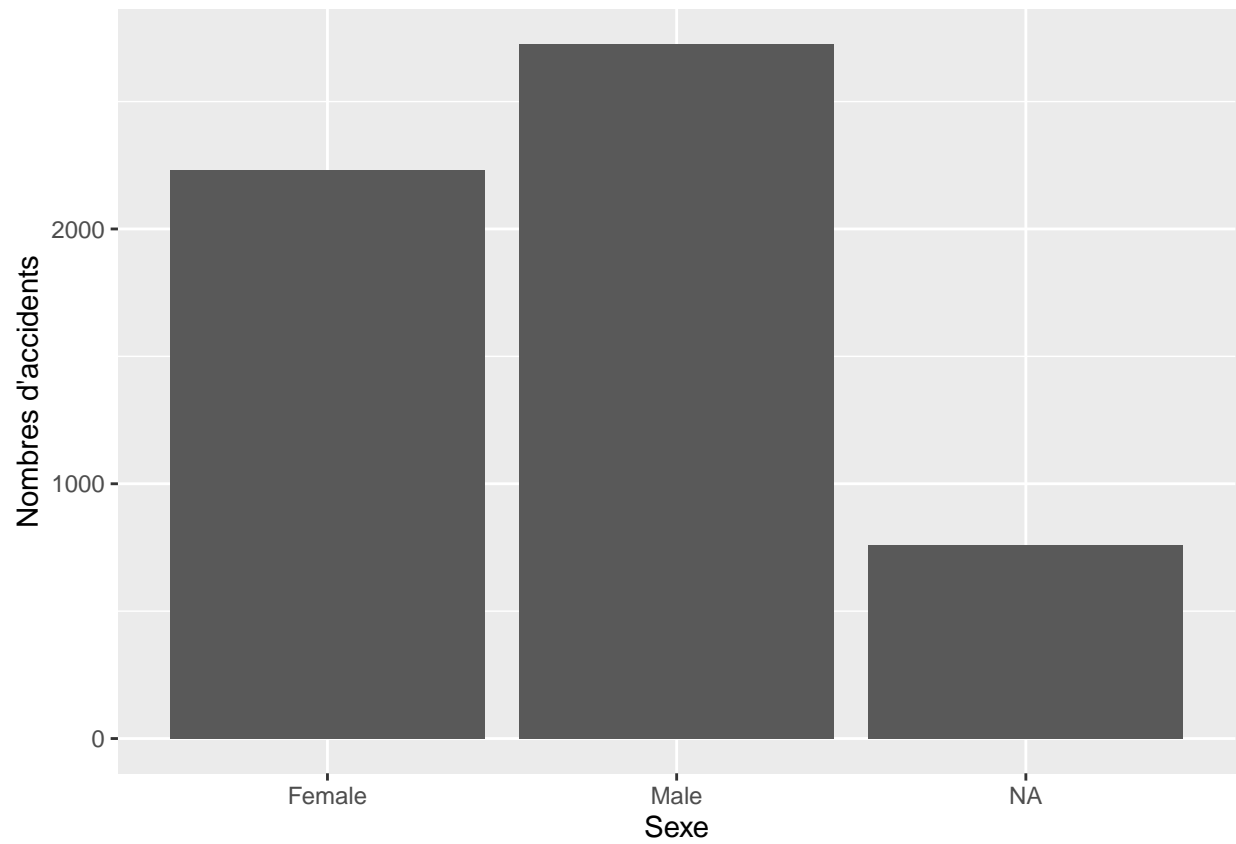
Var1	Freq
Female	2231
Male	2727

## Repartitions de la gravite des blessures par type de route et positionnement



## 2 Etude sur les conducteurs

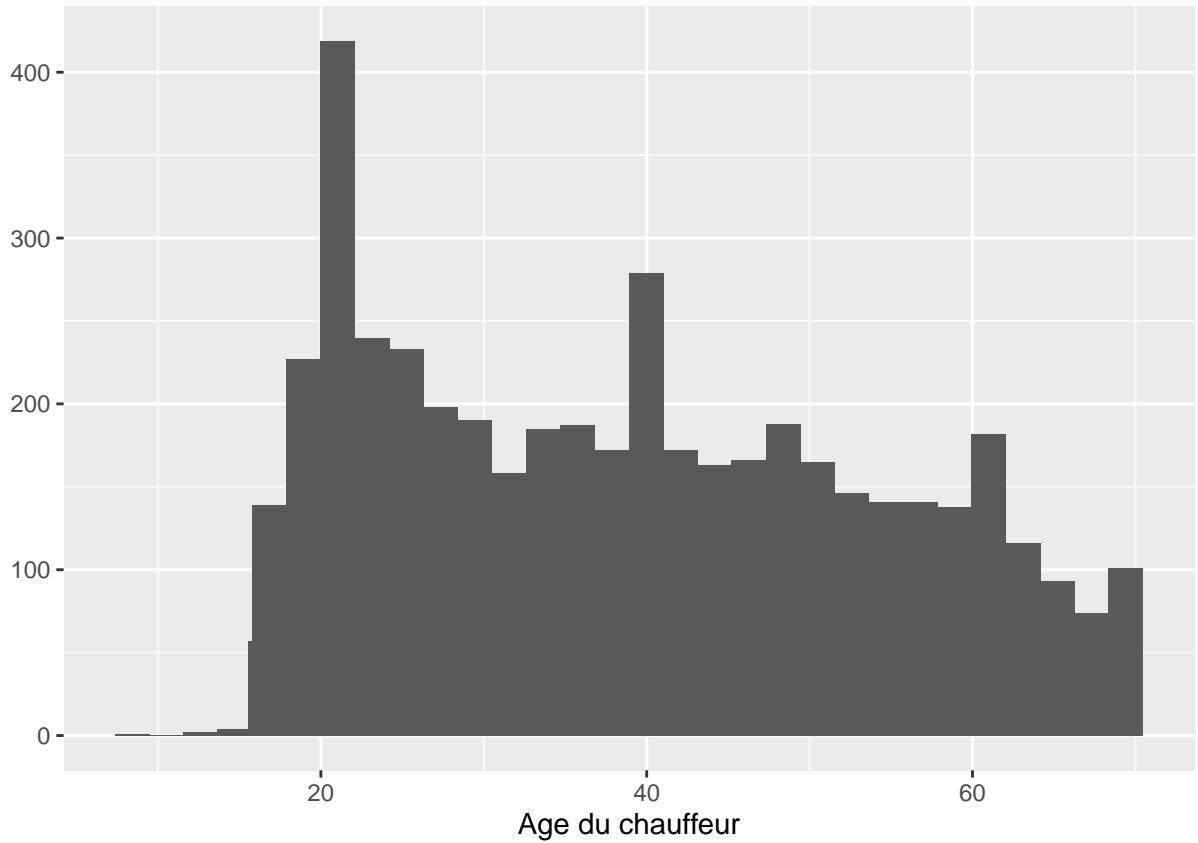
## 2.1 Sexe du conducteur



Contrairement aux accidentés qui sont essentiellement des hommes, le sexe des auteurs des accidents est quand à lui repartit plus quitablement entre les hommes et les femmes. Cependant le doute subsiste étant donné le nombre de données manquantes.

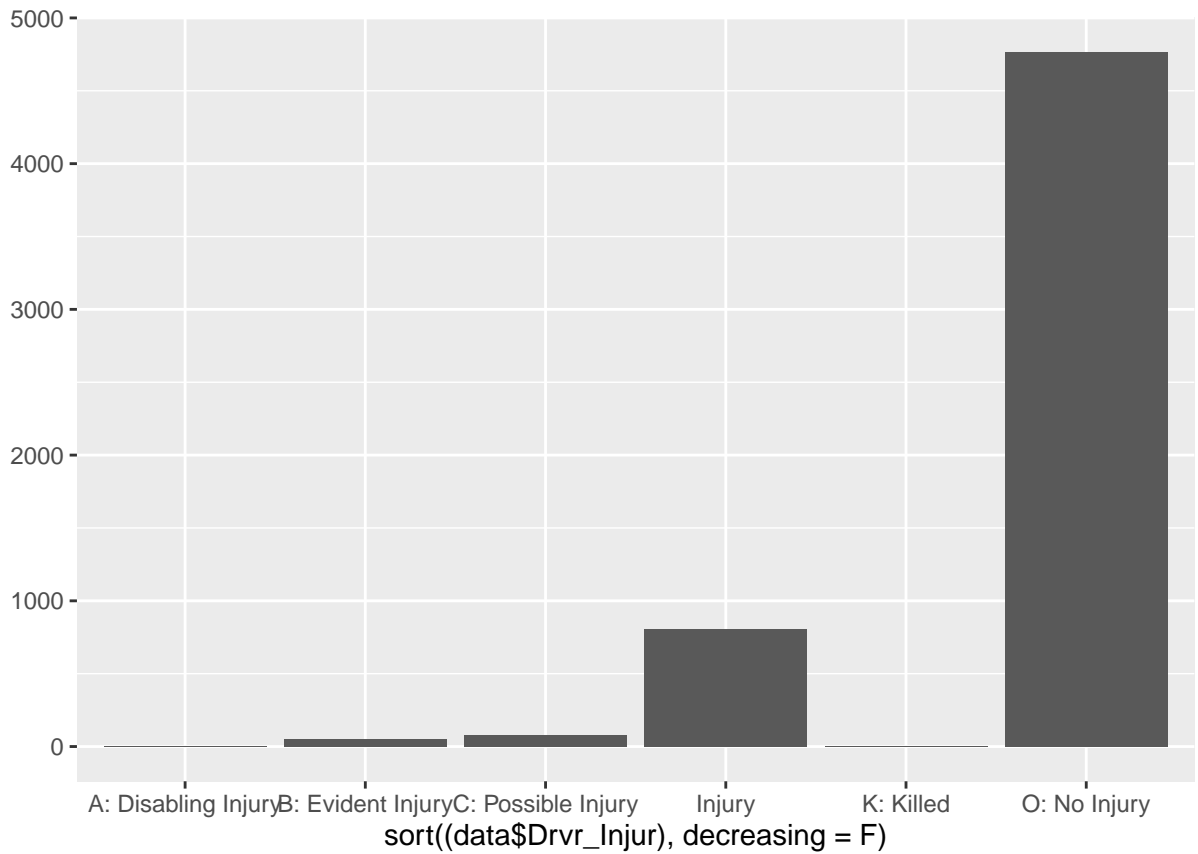
Var1	Freq
O: No Injury	0.8341498
Injury	0.1410077
C: Possible Injury	0.0143457
B: Evident Injury	0.0094472
A: Disabling Injury	0.0006998
K: Killed	0.0003499

## 2.2 Age des chauffeurs.



L'âge du chauffeur est la variable ayant le plus de données manquantes. Cette représentation graphique permet néanmoins d'observer la distribution de l'âge des chauffeurs. Trois pics sont a notés, le premier a 21 ans , le second à 40 et enfin le dernier a 60 ans.

## 2.3 Blessures des chauffeurs

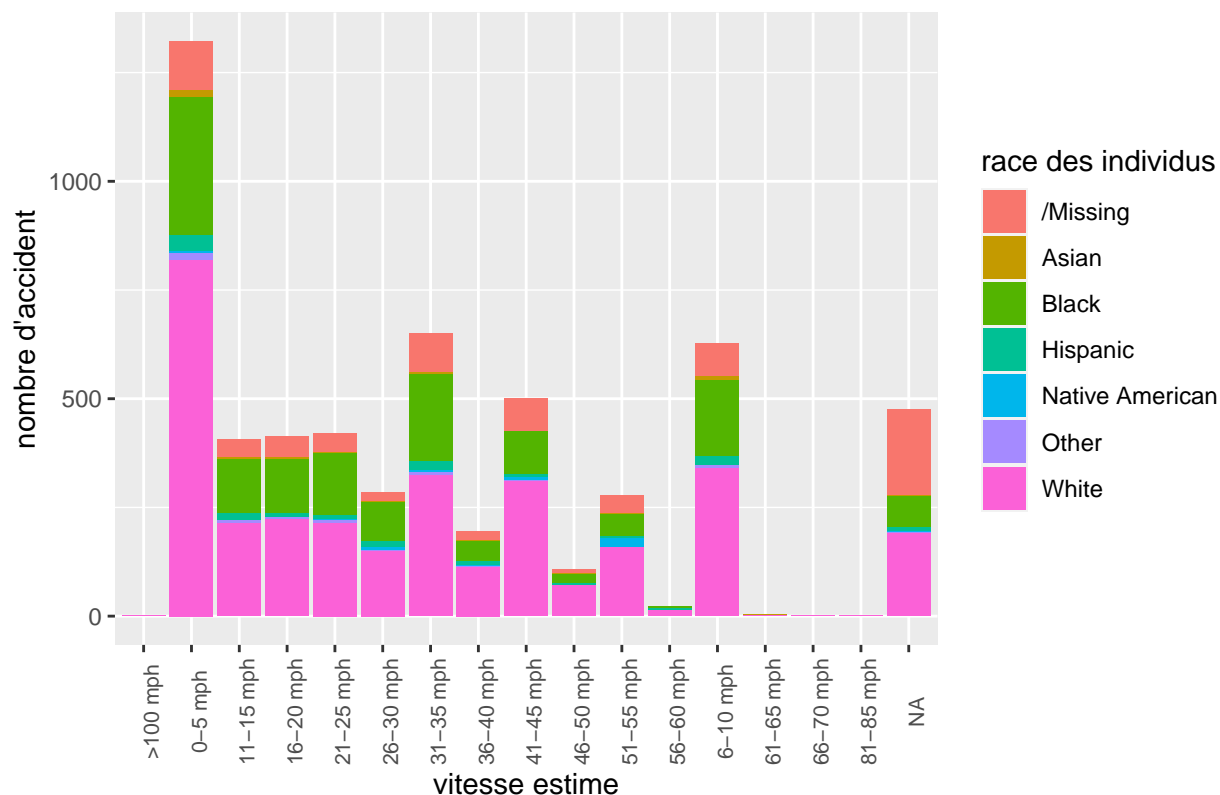


On aurait pu supposer d'instinct qu'un accident entre une voiture et un cycliste ne causera que très peu de dommage au conducteur de la voiture, les données le confirme. Dans la plupart des accidents le chauffeur s'en est sorti sans blessures. Il n'y a que 2 morts parmi les 5716 accidents recensés.

	/Missing	Asian	Black	Hispanic	Native American	Other	White
>100 mph	0	0	0	0	0	0	2
0-5 mph	114	14	317	39	5	14	819
11-15 mph	42	5	123	15	2	7	213
16-20 mph	48	5	123	9	2	5	222
21-25 mph	43	4	142	8	4	6	214
26-30 mph	22	1	91	15	5	2	150
31-35 mph	92	3	201	20	6	6	324
36-40 mph	21	3	44	10	2	2	113
41-45 mph	74	1	99	7	7	3	309
46-50 mph	8	1	23	3	0	2	70
51-55 mph	42	1	50	6	20	1	158
56-60 mph	0	0	4	5	0	0	13
6-10 mph	74	11	175	20	1	6	340
61-65 mph	1	1	1	0	0	0	2
66-70 mph	0	0	1	0	0	0	1
81-85 mph	0	0	0	0	0	0	1

## 2.4 Relation entre la vitesse du chauffeur, sa couleur de peau et le nombre d'accidents

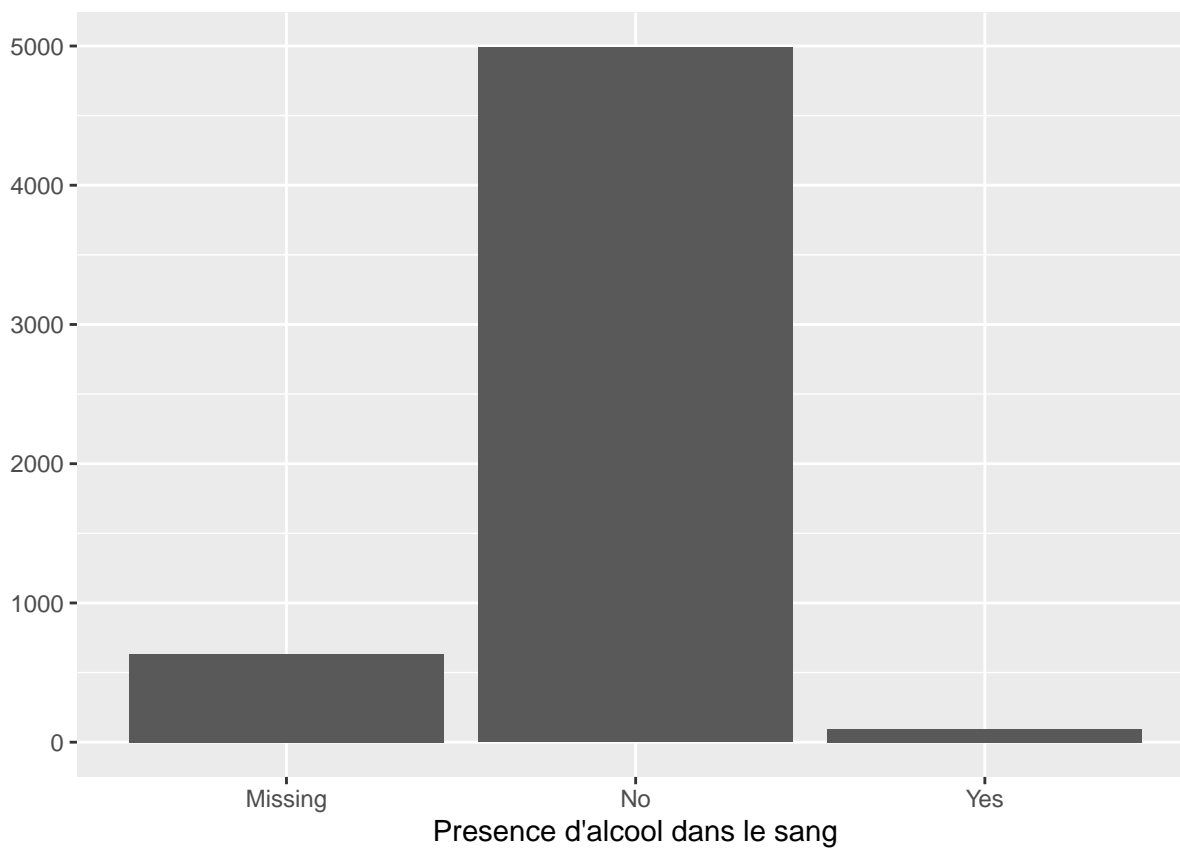
Histogramme des accidents par race et selon la vitesse



Les chauffeurs les plus impliqués dans un accidents sont de race Blanche, les second etant Noir. La plupart des individus impliqués dans un accident roulent entre 0 et 5 mph (milles par heure). Le fait que les victimes et les auteurs des accidents soient en majorité blanche peut s'expliquer simplement par le fait que la population soit en majorité blanche. Il serait intéressant de rapporter ces données a la proportion de la population, mais

nous manquons de données pour cela. Notons aussi que nous avons 1500 accidents dont nous n'avons pas de données sur la vitesse.

##Alcoolémie



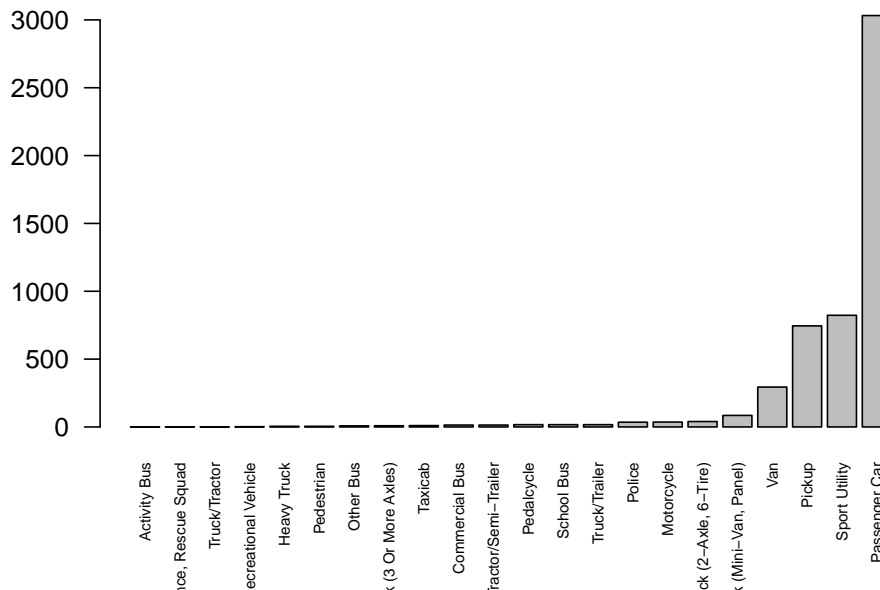
La plupart des chauffeurs n'étaient pas alcoolisés.

Var1	Freq
Missing	632
No	4990
Yes	94



Var1	Freq
Passenger Car	3032
Sport Utility	823
Pickup	745
Van	294
Light Truck (Mini-Van, Panel)	85
Single Unit Truck (2-Axle, 6-Tire)	40
Motorcycle	36
Police	35
Pedalcycle	17
School Bus	17
Truck/Trailer	17
Commercial Bus	14
Tractor/Semi-Trailer	14
Taxicab	10
Single Unit Truck (3 Or More Axles)	9
Other Bus	8
Heavy Truck	5
Pedestrian	5
Motor Home/Recreational Vehicle	2
Activity Bus	1
EMS Vehicle, Ambulance, Rescue Squad	1
Truck/Tractor	1

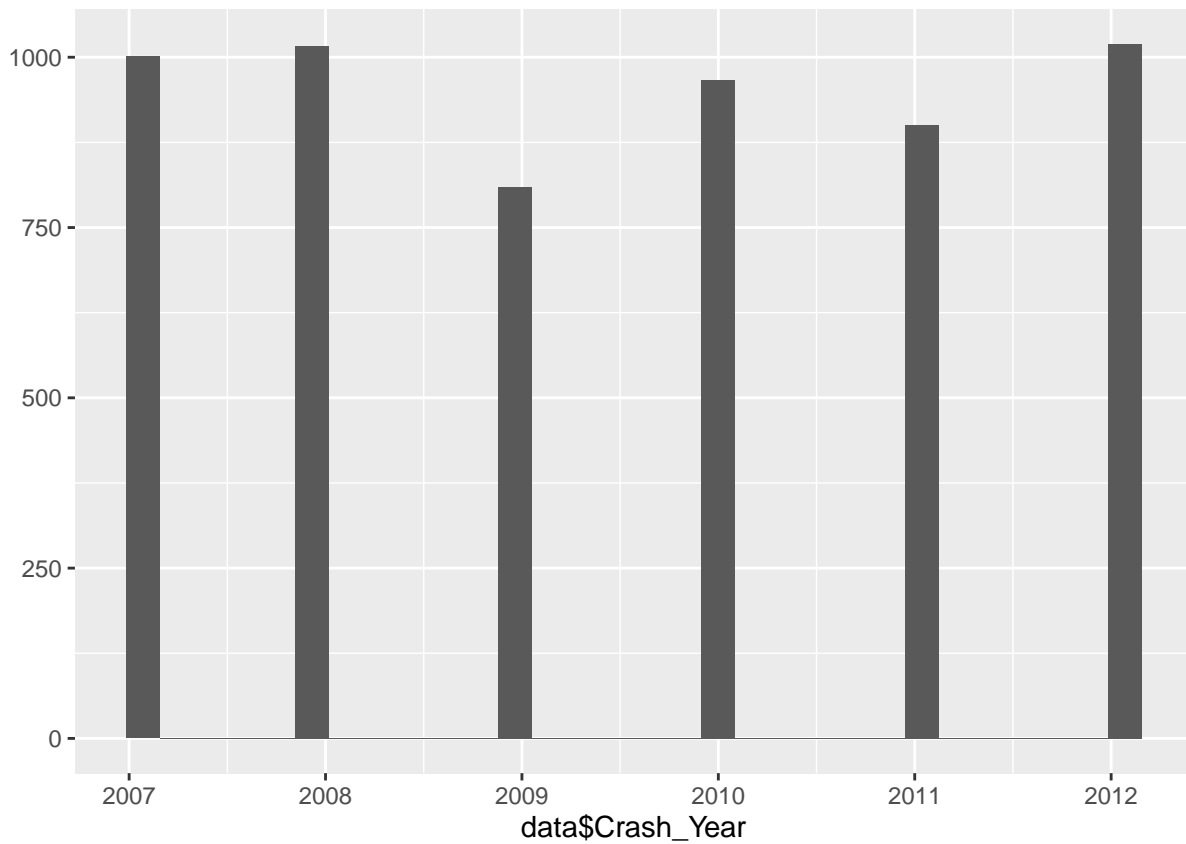
## 2.5 Type de voitures



Les principales types voitures impliqués dans des accidents de vélos sont des voitures familiales, voitures des port, Pickup et Van. Ceci pourrait s'expliquer par le fait que les cycliste se promenant surtout en agglomération ils rencontre surtout ce genre de voiture.

### 3 Etude sur les données temporelles

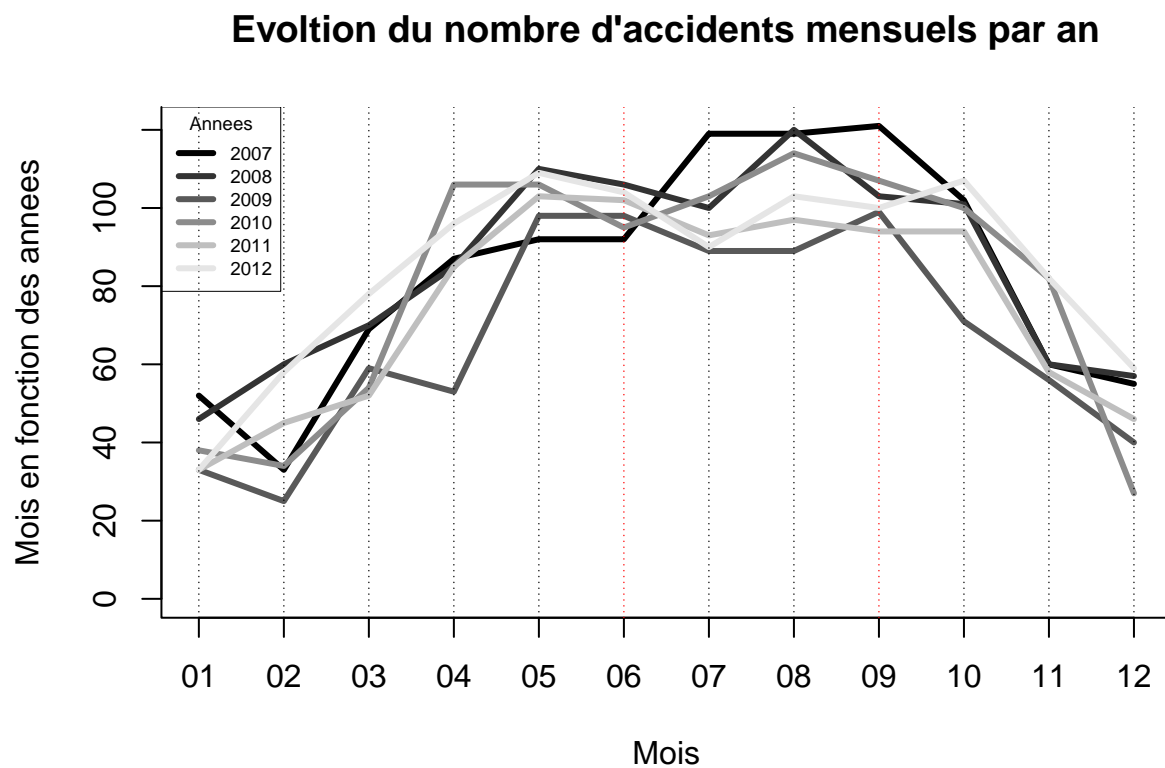
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Le premier graphique nous montre que globalement au fil des années le nombre d'accident impliquant un cycliste et un conducteur est resté plus ou moins le même au alentours de 1000 accident sauf au cours de l'année 2009 où nous avons eu une nette diminution d'environ 200 accidents.

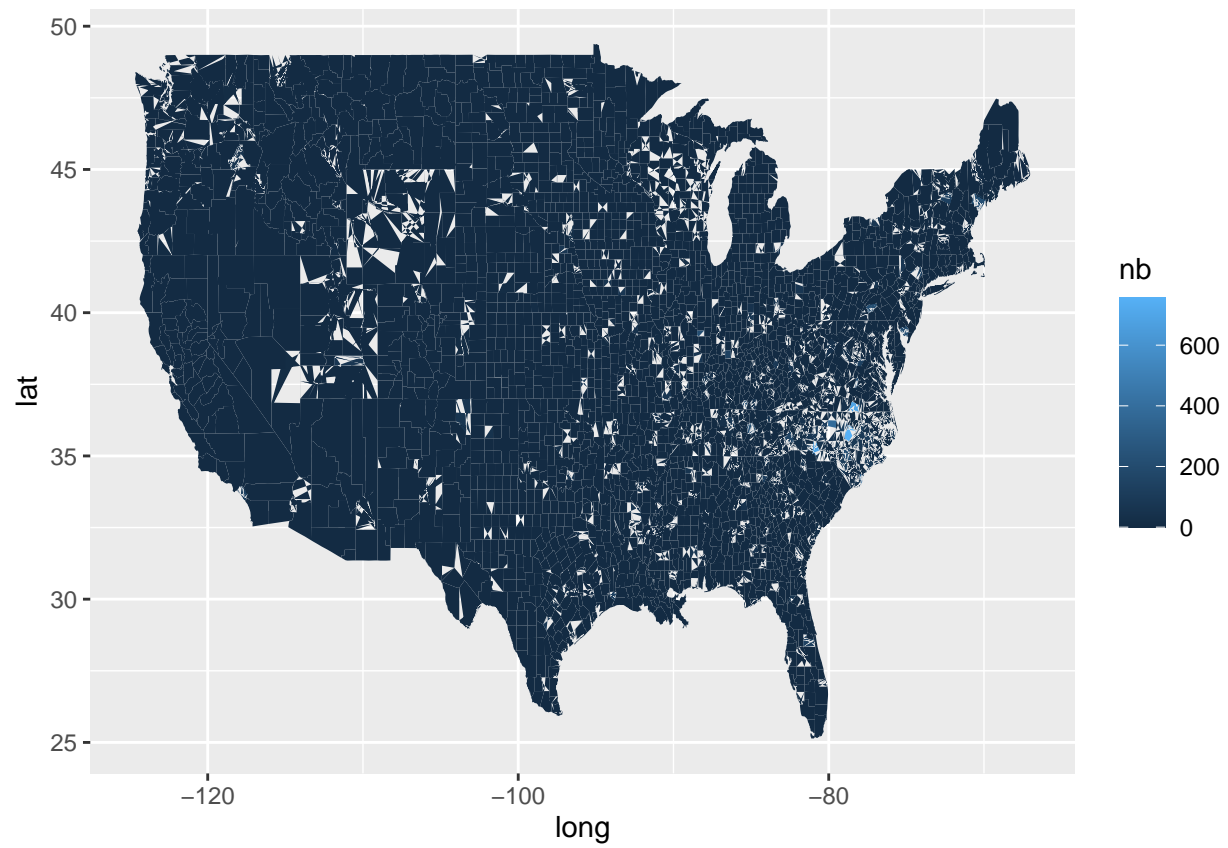
Nous avons des maintenant des informations concernant l'heure, le mois(extraite des la variable date) et l'année des differents accidents.

### 3.1 Evolution du Nombre d'accident par mois en fonction des années



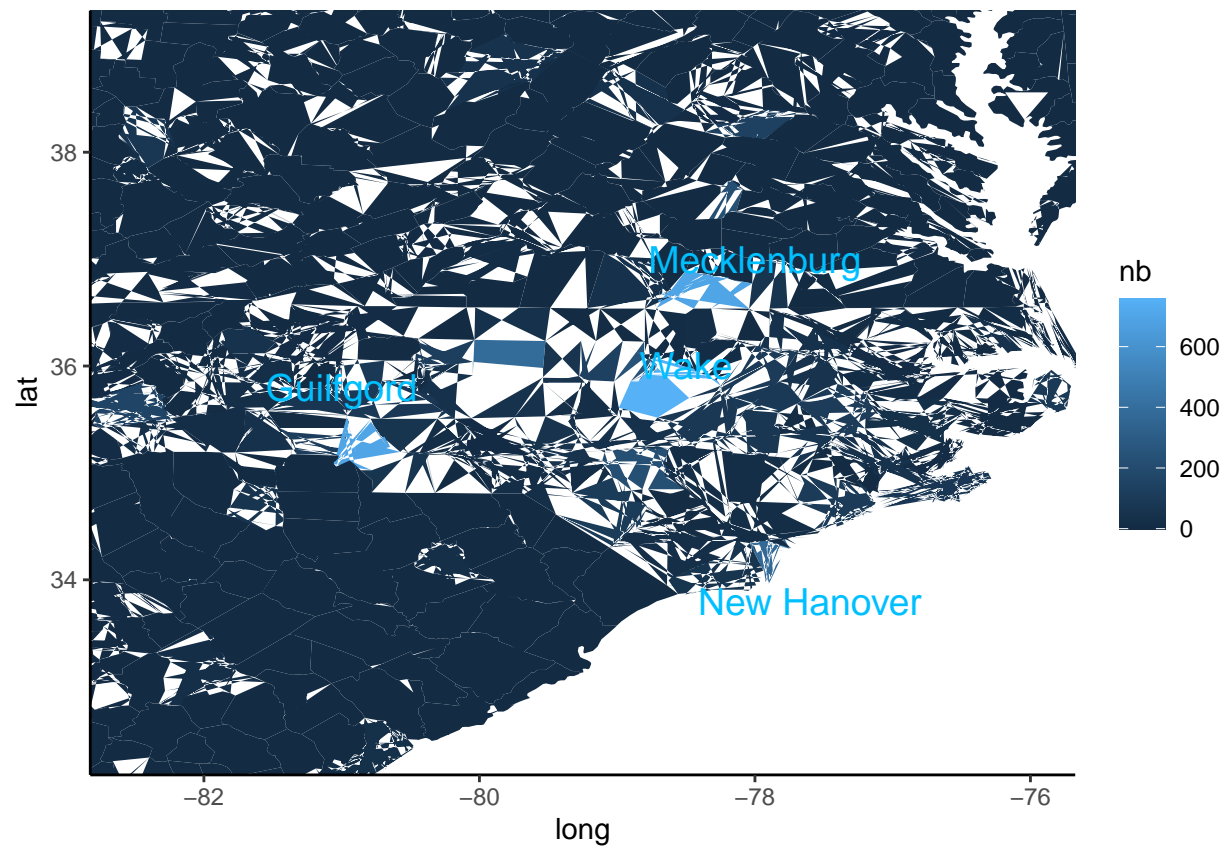
Les courbes representative de l'évolution des accidents en fonction des mois, par année , se chevauchant entre elles et étant données qu'aucune ne se démarque des autres, on peut supposer que le nombre d'accident est rester tendanciellennnt le même au cours de ces 5 années. On remarque aussi que pour pratiquement toutes les années le nombre d'accident est plus élevé entre le 6eme et le 9eme mois c'est a dire de juin à septembre.

## 4 Carte



### ## Zoom

Faisons un zoom sur les 4 comtés où le nombre d'accident est le plus important. Elles se trouve dans la meme zone géographique, l'est du Pays.



““