Dataset_Neiss

Arias Bonaud Peres

20/02/2020

Analyse des données de santé : épidémiologie et aide à la décision

Base de données d'interet (National Electronic Injury Surveillance System (NEISS))

La base de données collecte les cinq dernieres années du système de surveillance, (2013-2017) from the National Electronic Injury Surveillance System, which is a sample of all accidents reported to emergency rooms in the US. !(code) [https://github.com/hadley/neiss (https://github.com/hadley/neiss)] It currently contains three datasets:

```
injuries: individual injury results
products: product code lookup table
population: population of the US by age, sex, and year
```

```
# Package from dev version in github
# install.packages("devtools")
# devtools::install_github("hadley/neiss")
```

Lib in linux `sudo apt-get install libssl-dev libxml2-dev`

Load data

```
library("neiss")
data <- as.data.frame(injuries)
str(data)
```

```
## 'data.frame':    1865651 obs. of  18 variables:
##  $ case_num   : chr  "130104962" "130104963" "130104966" "1301
04968" ...
##  $ trmt_date  : Date, format: "2013-01-01" "2013-01-01" ...
##  $ age        : num  57 0.583 59 17 38 ...
##  $ sex        : chr  "Male" "Female" "Female" "Female" ...
##  $ race       : chr  "White" "Asian" "White" "White" ...
##  $ race_other : chr  NA NA NA NA ...
##  $ body_part  : chr  "Face" "Head" "Lower Trunk" "Ankle" ...
##  $ diag       : chr  "Contusion Or Abrasion" "Inter Organ Inju
ry" "Contusion Or Abrasion" "Strain, Sprain" ...
##  $ diag_other : chr  NA NA NA NA ...
##  $ disposition: chr  "Released" "Released" "Released" "Release
d" ...
##  $ location   : chr  "Sports Or Recreation Place" "Other Publi
c Property" "Home" "Home" ...
##  $ fmv        : chr  "No fire/flame/smoke" "No fire/flame/smok
e" "No fire/flame/smoke" "No fire/flame/smoke" ...
##  $ prod1      : num  3299 1807 1842 4076 474 ...
##  $ prod2      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ stratum    : chr  "M" "M" "M" "M" ...
##  $ psu        : num  100 100 100 100 100 100 100 94 61 61 ...
##  $ weight     : num  88.4 88.4 88.4 88.4 88.4 ...
##  $ narrative  : chr  "57YOM FELL WHILE JOGGING ON TRAIL DX: CO
NTUSION TO FACE" "7MOF HIT HEAD ON FLOOR AT DAY-CARE DX: CLOSED H
EAD INJURY" "59YOF FELL WHILE ON STAIRS DX: CONTUSION TO BUTTOCKS
" "17YOF TWISTED ANKLE STEPPING OUT OF BED DX: ANKLE STRAIN" ...
```

Transformation to categorical data

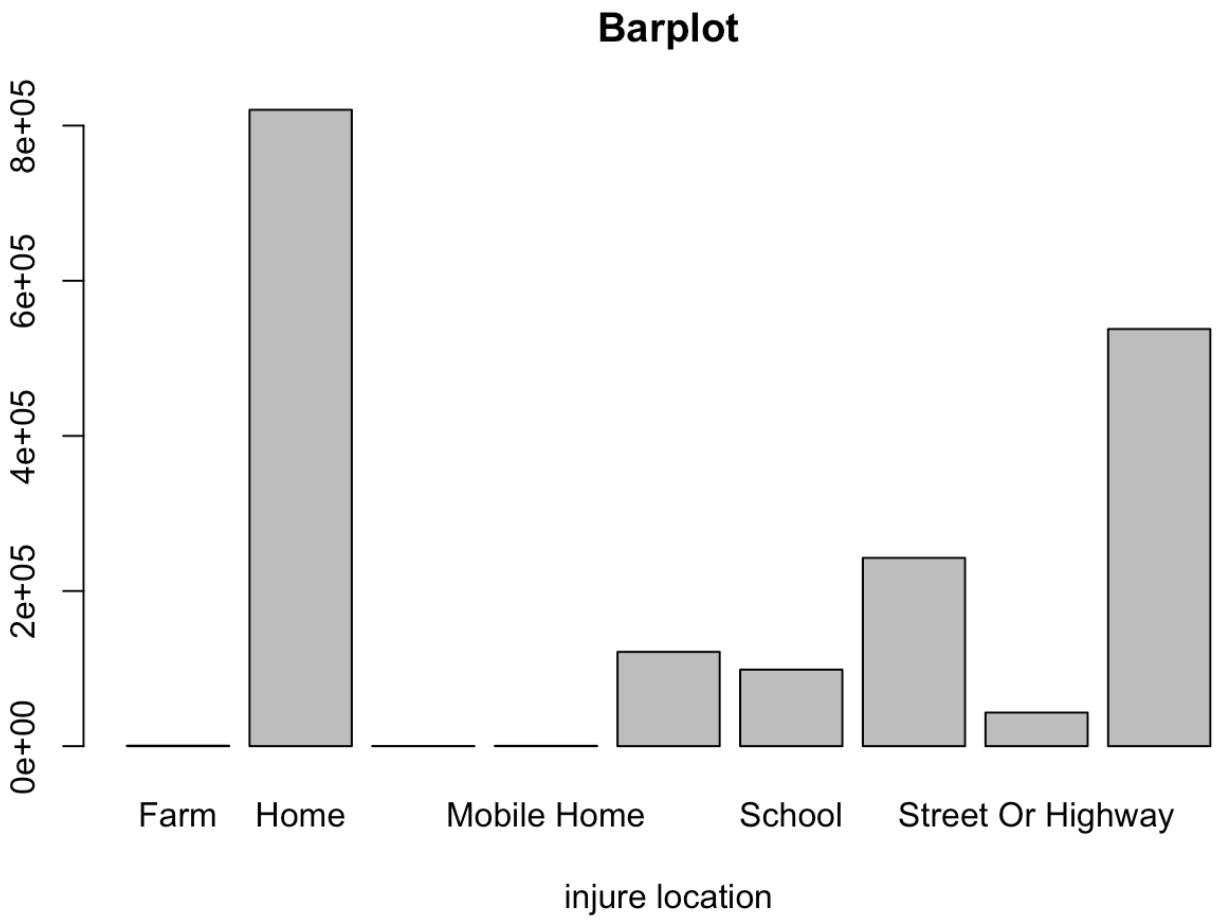
```
dim(data)
```

```
## [1] 1865651      18
```

```
data$sex <- as.factor(data$sex)
data$body_part <- as.factor(data$body_part)
data$location <- as.factor(data$location)
#change others columns ....
str(data)
```

```
## 'data.frame':    1865651 obs. of  18 variables:
##  $ case_num   : chr  "130104962" "130104963" "130104966" "1301
04968" ...
##  $ trmt_date  : Date, format: "2013-01-01" "2013-01-01" ...
##  $ age        : num  57 0.583 59 17 38 ...
##  $ sex        : Factor w/ 3 levels "Female","Male",..: 2 1 1 1
2 1 2 1 2 2 ...
##  $ race       : chr  "White" "Asian" "White" "White" ...
##  $ race_other : chr  NA NA NA NA ...
##  $ body_part  : Factor w/ 26 levels "25 - 50% Body",..: 7 11 1
6 3 8 26 2 9 13 3 ...
##  $ diag       : chr  "Contusion Or Abrasion" "Inter Organ Inju
ry" "Contusion Or Abrasion" "Strain, Sprain" ...
##  $ diag_other : chr  NA NA NA NA ...
##  $ disposition: chr  "Released" "Released" "Released" "Release
d" ...
##  $ location   : Factor w/ 9 levels "Farm","Home",..: 7 5 2 2 2
7 2 2 2 7 ...
##  $ fmv        : chr  "No fire/flame/smoke" "No fire/flame/smok
e" "No fire/flame/smoke" "No fire/flame/smoke" ...
##  $ prod1      : num  3299 1807 1842 4076 474 ...
##  $ prod2      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ stratum    : chr  "M" "M" "M" "M" ...
##  $ psu        : num  100 100 100 100 100 100 100 94 61 61 ...
##  $ weight     : num  88.4 88.4 88.4 88.4 88.4 ...
##  $ narrative  : chr  "57YOM FELL WHILE JOGGING ON TRAIL DX: CO
NTUSION TO FACE" "7MOF HIT HEAD ON FLOOR AT DAY-CARE DX: CLOSED H
EAD INJURY" "59YOF FELL WHILE ON STAIRS DX: CONTUSION TO BUTTOCKS
" "17YOF TWISTED ANKLE STEPPING OUT OF BED DX: ANKLE STRAIN" ...
```

```
plot(data$location,xlab = "injure location", main = "Barplot")
```



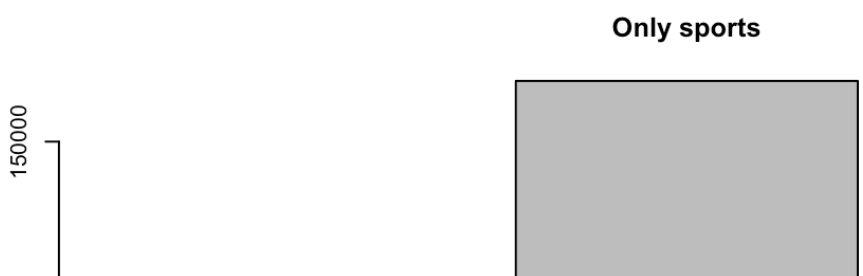
Just injuries from sport

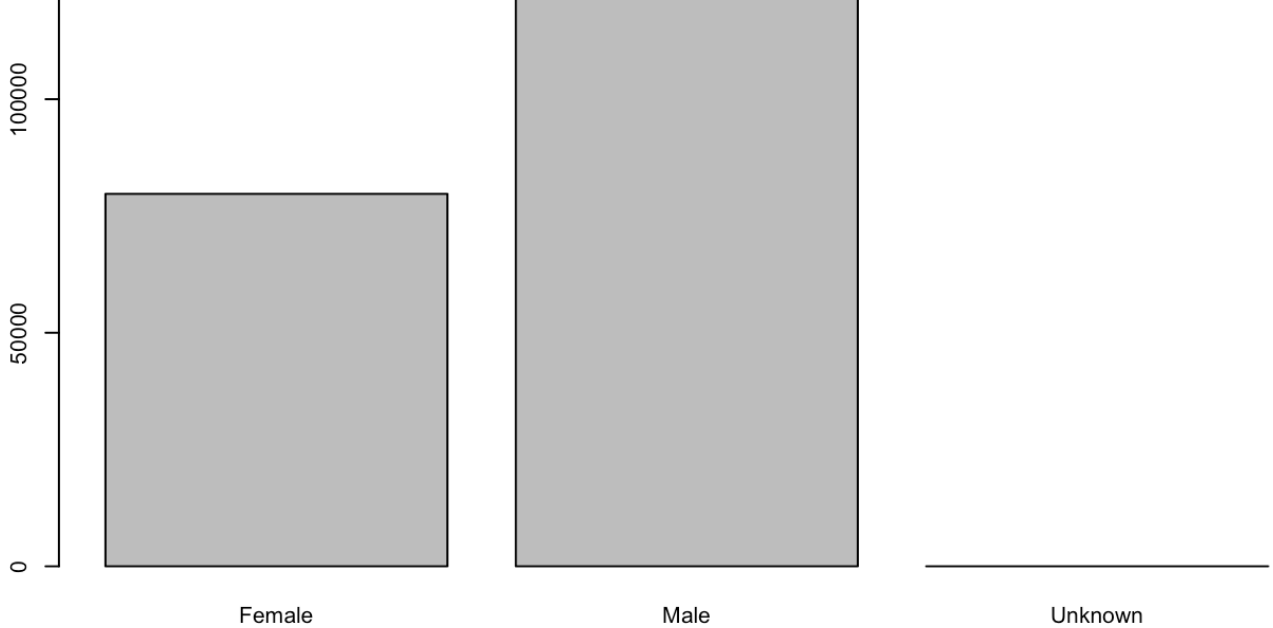
```
str(data$location)
```

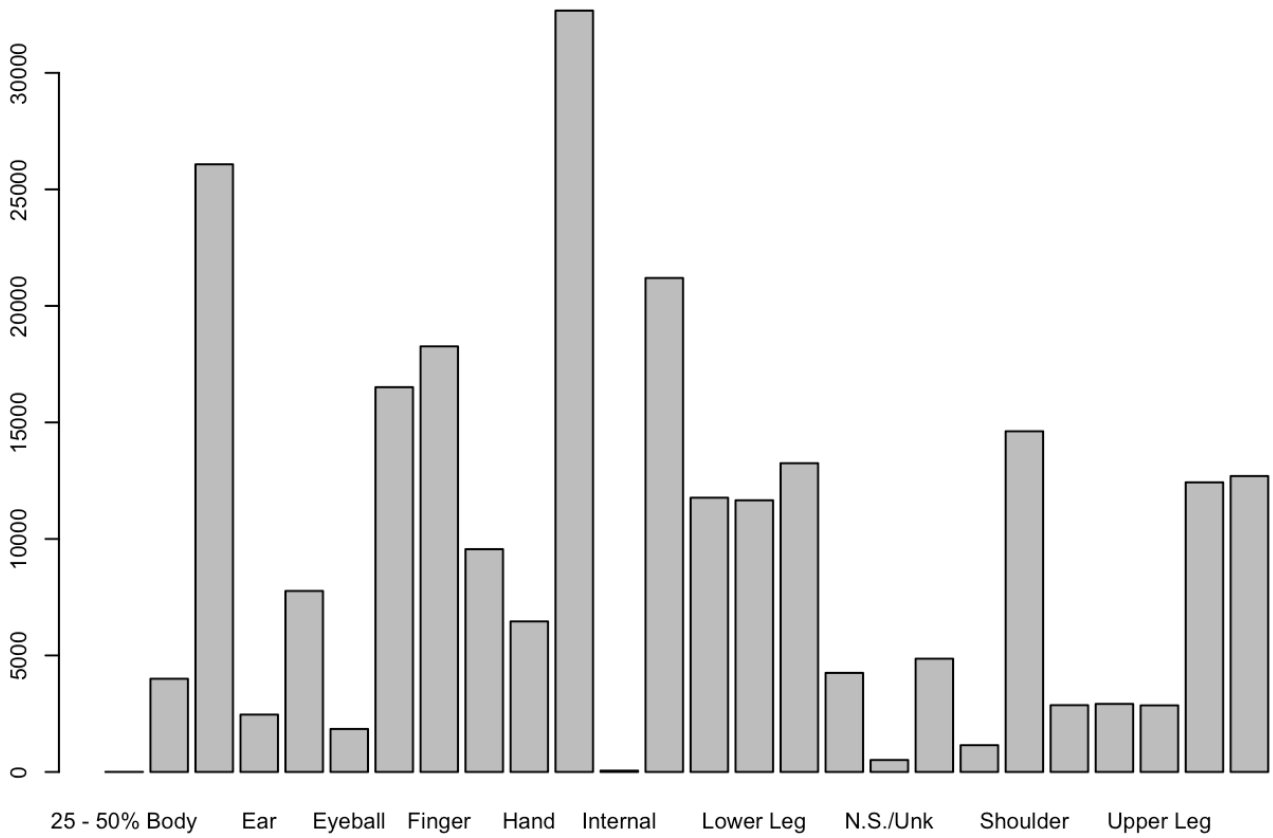
```
##  Factor w/ 9 levels "Farm","Home",..: 7 5 2 2 2 7 2 2 2 7 ...
```

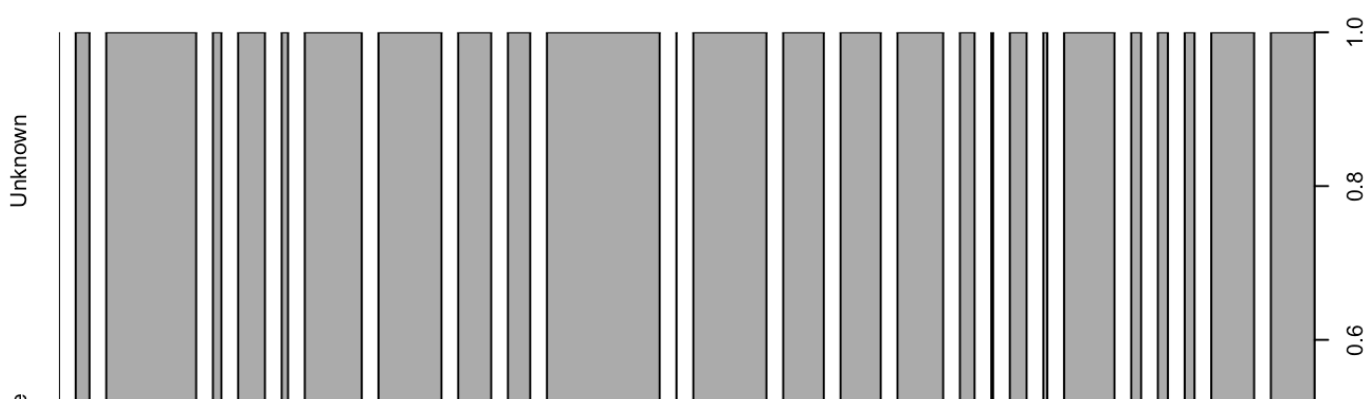
```
sport <- data[which(data$location == "Sports Or Recreation Place"
),]
```

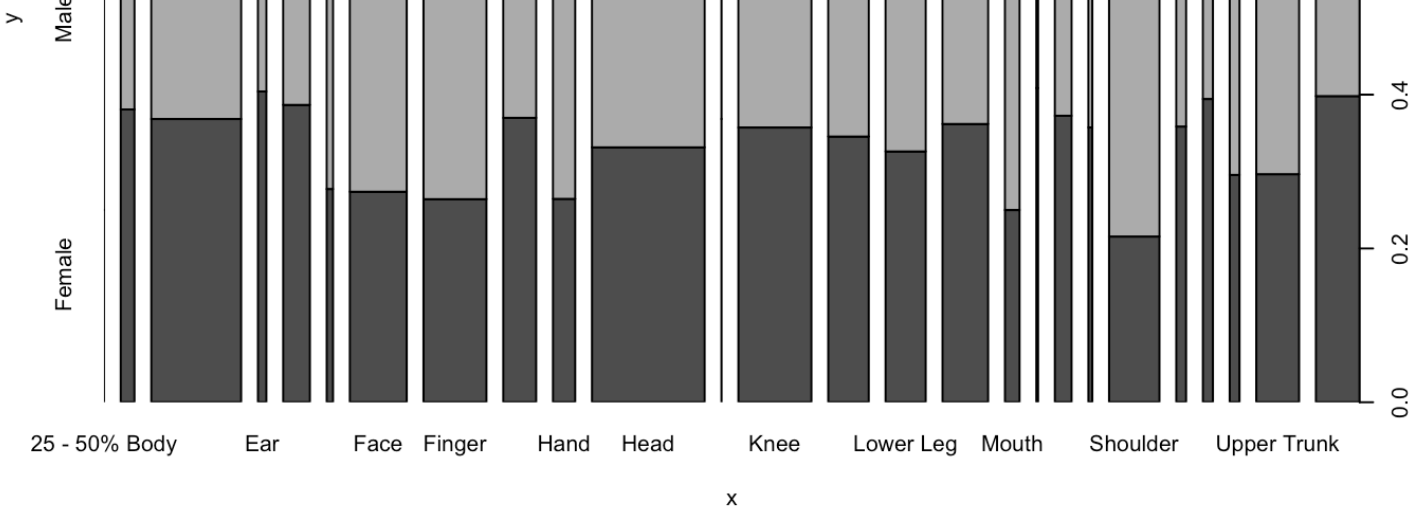
```
p <-par(mfrow=c(3,1))
plot(sport$sex, main ="Only sports")
plot(sport$body_part)
plot(sport$body_part,sport$sex)
```




0
50000
100000
Female
Male
Unknown


0
5000
10000
15000
20000
25000
30000
25 - 50% Body
Ear
Eyeball
Finger
Hand
Internal
Lower Leg
N.S./Unk
Shoulder
Upper Leg


Unknown
0.6
0.8
1.0


y
Male
Female
0.0
0.2
0.4
25 - 50% Body
Ear
Face
Finger
Hand
Head
Knee
Lower Leg
Mouth
Shoulder
Upper Trunk
x