

Statistiques population cible 2018

```
library(tidyverse)

## Error : 'format_warning' is not an exported object from 'namespace:cli'

## -- Attaching packages -----
## ----- tidyverse 1.2.1 -----

## v ggplot2 3.2.1      v purrr  0.3.3
## v tibble  2.1.3      v dplyr  0.8.3
## v tidyr   1.0.0      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.4.0

## -- Conflicts -----
## ----- tidyverse_conflicts() -----
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

DONNEES

Population totale

```
setwd("~/Documents/2A/Statapp/données")
library(readr)
nat2018us <- read_csv("W:/Documents/2A/Statapp/données/nat2018us.csv",
  col_types = cols_only(apgar5r = col_guess(),
                        dbwt =
col_guess(),
                        rf_inftr =
col_guess(),
                        rf_fedrg =
col_guess(),
                        rf_artec =
col_guess(),
                        mager =
col_guess(),
                        mrace6 =
col_guess(),
                        dmar =
col_guess(),
                        meduc =
col_guess(),
                        fagerec11 =
col_guess(),
                        frace6 =
col_guess(),
                        feduc =
```

```

col_guess(),
col_guess()))

priorlive =
dplural =

head.matrix(nat2018us)

## # A tibble: 6 x 14
##   mager mrace6 dmar meduc fagerec11 frace6 feduc priorlive rf_inftr
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <chr>
## 1    30     10     1     6         5     1     3         1 N
## 2    35     30     2     9         6     3     4         2 N
## 3    28     10     1     6         5     1     4         1 N
## 4    23     30     2     2         4     3     2         2 N
## 5    37     10     1     4         6     2     3         1 N
## 6    26     10     1     6         4     1     6         1 N
## # ... with 5 more variables: rf_fedrg <chr>, rf_artec <chr>,
## #   apgar5r <dbl>, dplural <dbl>, dbwt <dbl>

# nat2018us compte 2973577 observations

nat2018us <- subset(nat2018us,(nat2018us$mrace6 != 6 & nat2018us$meduc != 9 &
nat2018us$fagerec11 != 11 & nat2018us$frace6 != 9 & nat2018us$frace6 != 6 &
nat2018us$feduc != 9 & nat2018us$apgar5r != 5 & nat2018us$priorlive != 99))

```

Population cible

```

pop = subset(nat2018us,rf_inftr == "Y")
# il ne reste plus que 65921 observations

```

Quelques statistiques descriptives

```

table(nat2018us$rf_inftr)

##
##      N      U      Y
## 2906448  1208  65921

# 65921 ont recours à des traitements contre l'infertilité
round(table(nat2018us$rf_inftr)/length(nat2018us$rf_inftr),3)

##
##      N      U      Y
## 0.977 0.000 0.022

# 2.2%

# parmi ceux qui ont eu recours à des traitements contre l'infertilité
table(pop$rf_fedrg)

```

```
##
##      N      U      Y
## 34471  4137 27313

# 27313 ont recours à des traitements contre l'infertilité sous forme de
# médicaments ou insémination artificielle
round(table(pop$rf_fedrg)/length(pop$rf_fedrg),3) # soit environ 41%

##
##      N      U      Y
## 0.523 0.063 0.414

table(pop$rf_artec)

##
##      N      U      Y
## 22195  4137 39589

# 39589 sous forme d'assisted reproductive technology (ART) (IVF ou autres)
round(table(pop$rf_artec)/length(pop$rf_artec),3) # soit 60%

##
##      N      U      Y
## 0.337 0.063 0.601

# nombre d'enfant avant cette naissance
round(table(pop$priorlive)/length(pop$priorlive),3)

##
##      0      1      2      3      4      5      6      7      8      9     10     11
## 0.564 0.305 0.088 0.028 0.010 0.003 0.001 0.001 0.000 0.000 0.000 0.000
##      12
## 0.000

# De plus, 56% n'ont pas eu d'enfant auparavant
```

CARACTERISTIQUES DE LA POPULATION CIBLE

Age des parents

Age de la mère : 12 10 – 12 years 13 13 years ... 49 49 years 50 50 years and over

Age du père : 01 Under 15 years 02 15-19 years 03 20-24 years 04 25-29 years 05 30-34 years 06 35-39 years 07 40-44 years 08 45-49 years 09 50-54 years 10 55-98 years 11 Not stated

```
summary(select(pop,mager))
```

```
##      mager
## Min.    :16.0
## 1st Qu.:31.0
## Median :34.0
## Mean   :34.4
```

```
## 3rd Qu.:38.0
## Max.    :50.0

summary(select(pop,fagerec11))

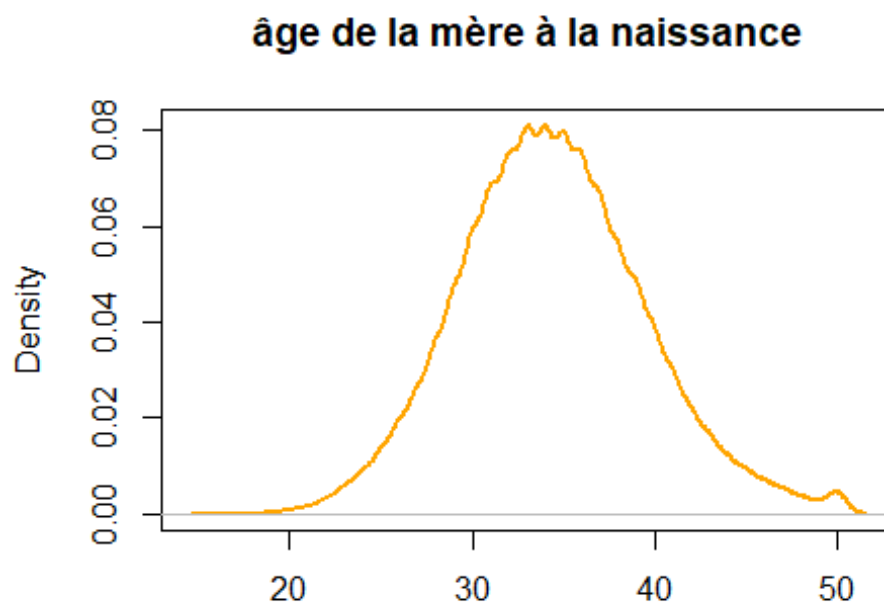
##      fagerec11
## Min.   : 2.00
## 1st Qu.: 5.00
## Median : 6.00
## Mean   : 5.93
## 3rd Qu.: 7.00
## Max.   :10.00

round(table(pop$fagerec11)/length(pop$fagerec11),3)

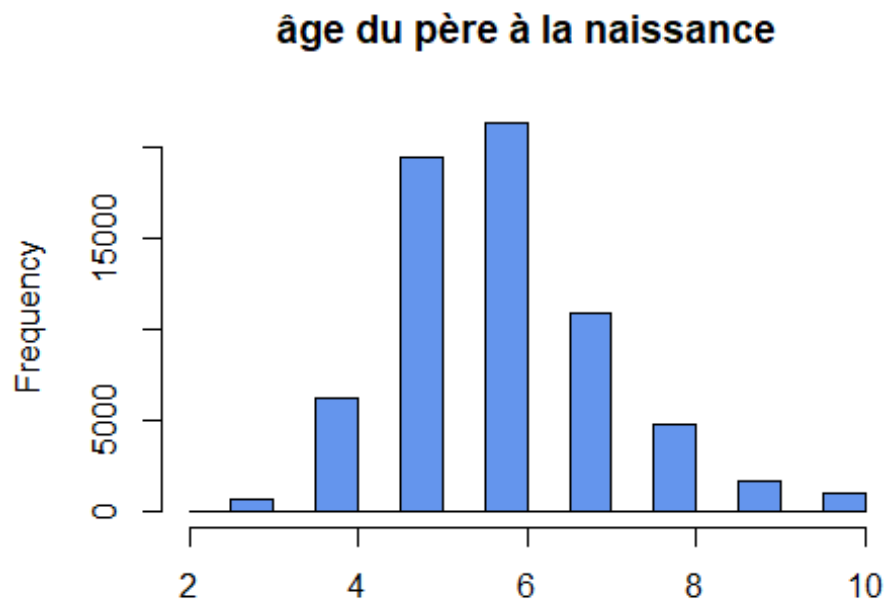
##
##      2      3      4      5      6      7      8      9     10
## 0.000 0.010 0.094 0.294 0.324 0.166 0.072 0.025 0.014

# 60% des pères sont âgés de 30 à 39 ans

plot(density(pop$mager), lwd = 2, col = "orange", xlab = "", main = "âge de
la mère à la naissance")
```



```
hist(pop$fagerec11, col = "cornflowerblue", xlab = "", main = "âge du père à
la naissance")
```



Les parents sont plus jeunes à la naissance de l'enfant sur la population entière

```
summary(select(nat2018us,mager))
```

```
##      mager
##  Min.   :13.00
## 1st Qu.:25.00
##  Median :30.00
##   Mean  :29.47
## 3rd Qu.:33.00
##   Max.  :50.00
```

```
summary(select(nat2018us,fagerec11))
```

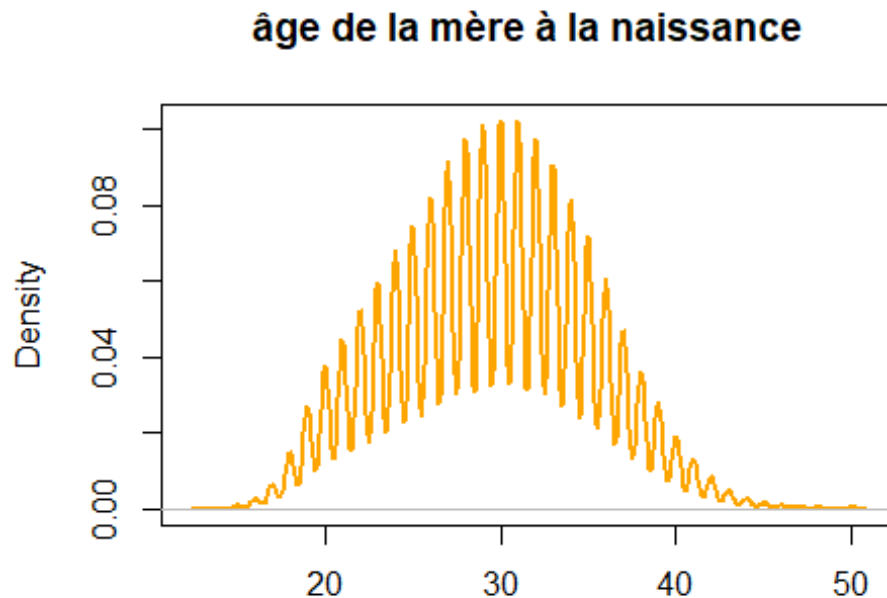
```
##      fagerec11
##  Min.   : 1.000
## 1st Qu.: 4.000
##  Median : 5.000
##   Mean  : 4.972
## 3rd Qu.: 6.000
##   Max.  :10.000
```

```
round(table(nat2018us$fagerec11)/length(nat2018us$fagerec11),3)
```

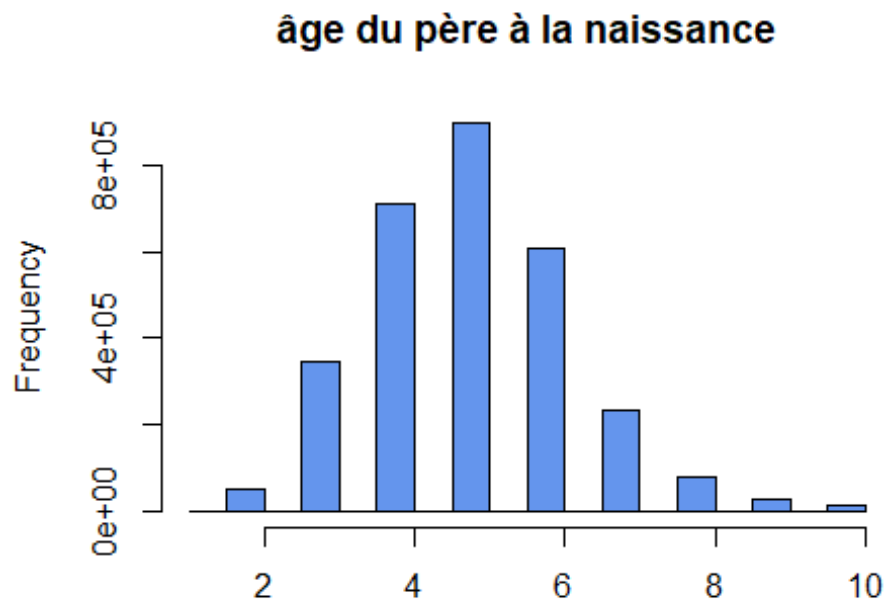
```
##
##      1      2      3      4      5      6      7      8      9     10
## 0.000 0.018 0.117 0.240 0.302 0.205 0.079 0.027 0.009 0.004
```

*# environ 55% des pères sont âgés de 25 à 34 ans
Alors que sur la population entière les pères sont âgés de 30 à 34 ans en moyenne, sur la population cible ils sont âgés de 35 à 39 ans en moyenne
L'écart est visible aussi pour les mères : 29-30 ans en moyenne / 34-35 ans chez la population cible*

```
plot(density(nat2018us$mager), lwd = 2, col = "orange", xlab = "", main =  
"âge de la mère à la naissance")
```



```
hist(nat2018us$fagerec11, col = "cornflowerblue", xlab = "", main = "âge du  
père à la naissance")
```



Ethnie des parents

Race de la mère : 1 White (only) 3 AIAN (American Indian or Alaskan Native) (only) 4 Asian (only) 5 NHOPI (Native Hawaiian or Other Pacific Islander) (only) 6 More than one race

Ethnie du père : 1.34, 75% blanc 1 White (only) 2 Black (only) 3 AIAN (only) 4 Asian (only) 5 NHOPI (only) 6 More than one race 9 Unknown or Not Stated

```
summary(select(pop,mrace6))
```

```
##      mrace6
##  Min.   :10.00
## 1st Qu.:10.00
##  Median :10.00
##   Mean   :15.03
## 3rd Qu.:10.00
##   Max.   :61.00
```

```
summary(select(pop,frace6))
```

```
##      frace6
##  Min.   :1.000
## 1st Qu.:1.000
##  Median :1.000
##   Mean   :1.392
## 3rd Qu.:1.000
##   Max.   :5.000
```

```

round(table(pop$frace6)/length(pop$frace6),3)

##
##      1      2      3      4      5
## 0.829 0.060 0.002 0.107 0.001

# 83 % des pères sont blancs, 11% asiatiques et 6% noirs
round(table(pop$mrace6)/length(pop$mrace6),3)

##
##     10     20     30     40     41     51     61
## 0.808 0.054 0.002 0.106 0.013 0.001 0.016

# taux similaires pour les mères

hist(pop$frace6, col = "cornflowerblue", xlab = "", main = "ethnie du père")

```



La part des pères blancs est toujours prédominante si on considère toute la population, mais elle l'est moins que chez la population cible et surtout la part des pères noirs est bien plus importante

```

summary(select(nat2018us, mrace6))

##      mrace6
## Min.      :10.0
## 1st Qu.:10.0
## Median :10.0
## Mean    :14.9

```



```
## 3rd Qu.:10.0
## Max. :61.0

summary(select(nat2018us,frace6))

##      frace6
## Min.   :1.000
## 1st Qu.:1.000
## Median :1.000
## Mean   :1.393
## 3rd Qu.:1.000
## Max.   :5.000

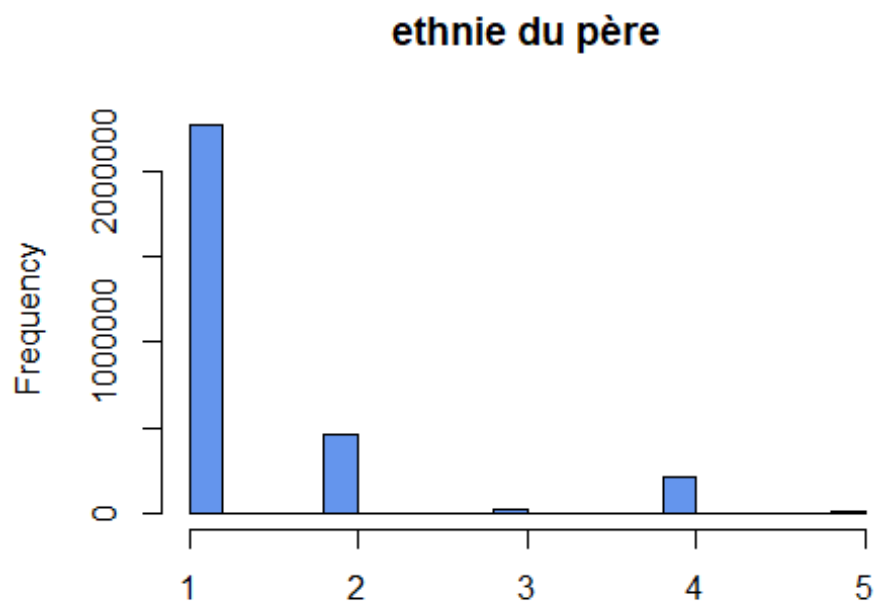
round(table(nat2018us$frace6)/length(nat2018us$frace6),3)

##
##      1      2      3      4      5
## 0.764 0.155 0.008 0.070 0.003

# 77 % des pères sont blancs, 16% noirs et 7% asiatiques
round(table(nat2018us$mrace6)/length(nat2018us$mrace6),3)

##
##     10     20     30     40     41     51     61
## 0.763 0.130 0.008 0.062 0.014 0.003 0.020

hist(nat2018us$frace6, col = "cornflowerblue", xlab = "", main = "ethnie du
père")
```



Education des parents éducation de la mère : 1 8th grade or less 2 9th through 12th grade with no diploma 3 High school graduate or GED completed 4 Some college credit, but not a degree. 5 Associate degree (AA,AS) 6 Bachelor's degree (BA, AB, BS) 7 Master's degree (MA, MS, MEng, MEd, MSW, MBA) 8 Doctorate (PhD, EdD) or Professional Degree (MD, DDS,DVM, LLB, JD) 9 Unknown

Education du père : 5.6, 1 8th grade or less 2 9th through 12th grade with no diploma 3 High school graduate or GED completed 4 Some college credit, but not a degree. 5 Associate degree (AA,AS) 6 Bachelor's degree (BA, AB, BS) 7 Master's degree (MA, MS, MEng, MEd, MSW, MBA) 8 Doctorate (PhD, EdD) or Professional Degree (MD, DDS,DVM, LLB, JD) 9 Unknown

```
summary(select(pop,meduc))
```

```
##      meduc
##  Min.    :1.000
## 1st Qu.:5.000
##  Median :6.000
##   Mean   :5.902
## 3rd Qu.:7.000
##   Max.   :8.000
```

```
summary(select(pop,feduc))
```

```
##      feduc
##  Min.    :1.000
## 1st Qu.:4.000
##  Median :6.000
##   Mean   :5.612
## 3rd Qu.:7.000
##   Max.   :8.000
```

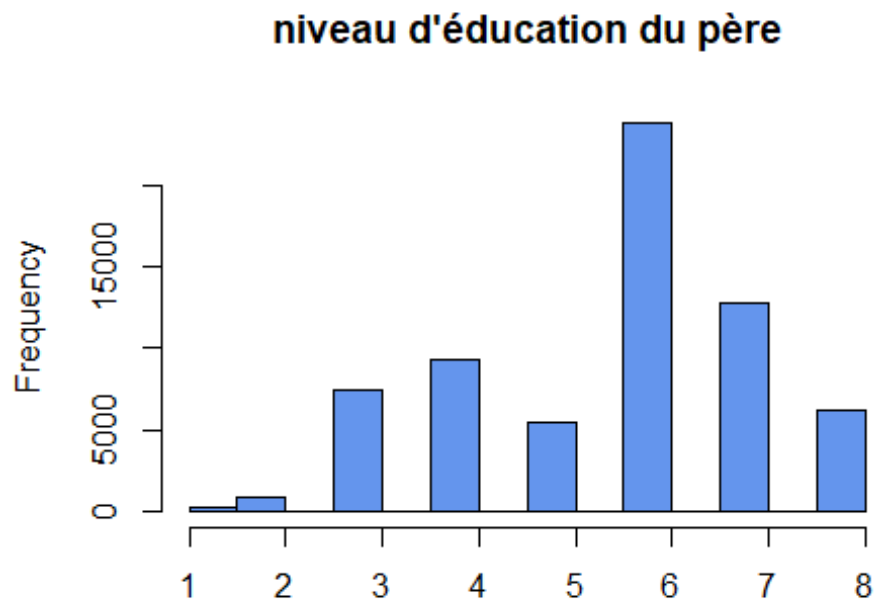
```
round(table(pop$meduc)/length(pop$meduc),3)
```

```
##
##      1      2      3      4      5      6      7      8
## 0.003 0.008 0.066 0.109 0.083 0.375 0.261 0.094
```

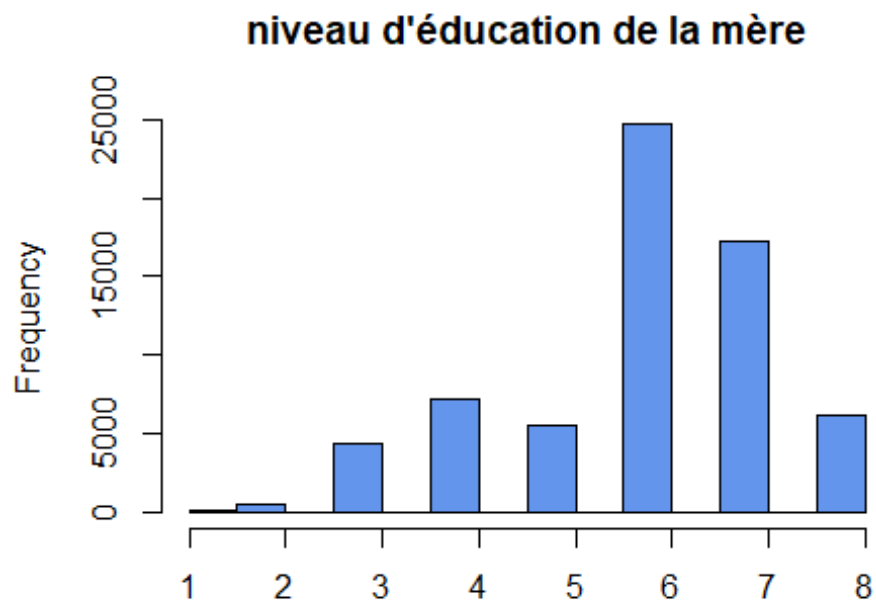
```
round(table(pop$feduc)/length(pop$feduc),3)
```

```
##
##      1      2      3      4      5      6      7      8
## 0.004 0.012 0.112 0.141 0.082 0.361 0.193 0.095
```

```
hist(pop$feduc, col = "cornflowerblue", xlab = "", main = "niveau
d'éducation du père")
```



```
hist(pop$meduc, col = "cornflowerblue", xlab = "", main = "niveau  
d'éducation de la mère")
```



Sur la population entière, le niveau d'éducation des parents est bien moins élevé : en moyenne les parents qui ont recours au traitement ont un Bachelor's degree, alors que toute la population considérée, ils ont "Some college credit, but not a degree". L'écart se voit en particulier à la part de haut niveau d'éducation : plus de 20% des parents de la population cible ont un master ou un doctorat, tandis que sont plutôt 11% des parents en général.

```
summary(select(nat2018us,meduc))
```

```
##      meduc
##  Min.    :1.000
## 1st Qu.:3.000
##  Median :4.000
##   Mean   :4.597
## 3rd Qu.:6.000
##   Max.   :8.000
```

```
summary(select(nat2018us,feduc))
```

```
##      feduc
##  Min.    :1.000
## 1st Qu.:3.000
##  Median :4.000
##   Mean   :4.342
## 3rd Qu.:6.000
##   Max.   :8.000
```

```
round(table(nat2018us$meduc)/length(nat2018us$meduc),3)
```

```
##
##      1      2      3      4      5      6      7      8
## 0.024 0.071 0.231 0.197 0.090 0.240 0.113 0.033
```

```
round(table(nat2018us$feduc)/length(nat2018us$feduc),3)
```

```
##
##      1      2      3      4      5      6      7      8
## 0.029 0.082 0.296 0.189 0.075 0.209 0.083 0.037
```

Caractéristiques du couple

Statut marital : 1 Married 2 Unmarried

Enfant(s) encore en vie : 00-30 Number of children still living from previous live births. 99 Unknown or not stated

```
summary(select(pop,dmar))
```

```
##      dmar
##  Min.    :1.000
## 1st Qu.:1.000
##  Median :1.000
```

```
## Mean :1.041
## 3rd Qu.:1.000
## Max. :2.000
## NA's :5844

summary(select(pop,priorlive))

## priorlive
## Min. : 0.0000
## 1st Qu.: 0.0000
## Median : 0.0000
## Mean : 0.6384
## 3rd Qu.: 1.0000
## Max. :12.0000

summary(select(pop,dplural))

## dplural
## Min. :1.000
## 1st Qu.:1.000
## Median :1.000
## Mean :1.246
## 3rd Qu.:1.000
## Max. :5.000

round(table(pop$dmar)/length(pop$dmar),3) # 87% sont mariés

##
## 1 2
## 0.874 0.037

round(table(pop$priorlive)/length(pop$priorlive),3) # 86% ont déjà au plus un
enfant

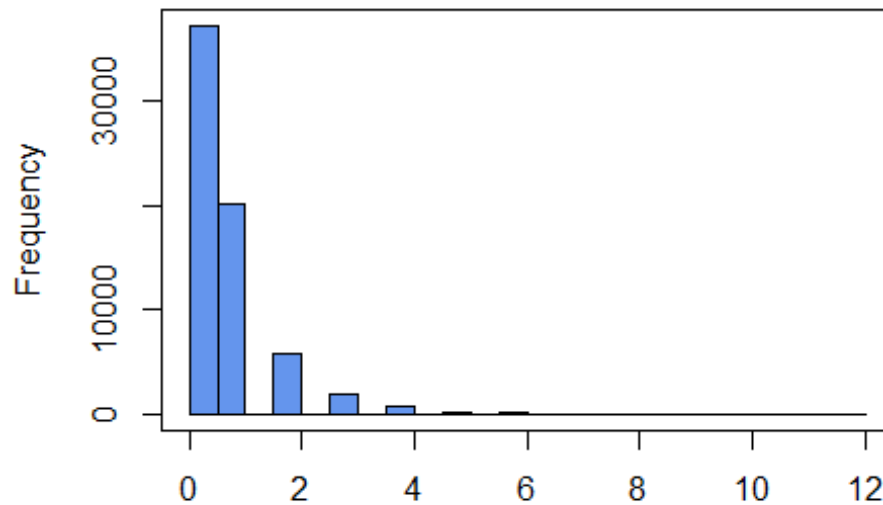
##
## 0 1 2 3 4 5 6 7 8 9 10 11
## 0.564 0.305 0.088 0.028 0.010 0.003 0.001 0.001 0.000 0.000 0.000 0.000
## 12
## 0.000

round(table(pop$dplural)/length(pop$dplural),3) # 20% des mères ont donné
naissance à des jumeaux

##
## 1 2 3 4 5
## 0.771 0.213 0.015 0.001 0.000

hist(pop$priorlive, col = "cornflowerblue", xlab = "", main = "nombre
d'enfants")
box()
```

nombre d'enfants



Moins de couples sont mariés et plus ont déjà eu des enfants auparavant mais surtout, beaucoup moins de naissance de jumeaux

```
summary(select(nat2018us,dmar))
```

```
##      dmar
##  Min.   :1.0
## 1st Qu.:1.0
##  Median :1.0
##   Mean  :1.3
## 3rd Qu.:2.0
##   Max.  :2.0
## NA's   :363340
```

```
summary(select(nat2018us,priorlive))
```

```
##    priorlive
##  Min.   : 0.000
## 1st Qu.: 0.000
##  Median : 1.000
##   Mean  : 1.102
## 3rd Qu.: 2.000
##   Max.  :18.000
```

```
summary(select(nat2018us,dplural))
```

```
##      dplural
##  Min.   :1.000
```

```
## 1st Qu.:1.000
## Median :1.000
## Mean :1.035
## 3rd Qu.:1.000
## Max. :5.000

round(table(nat2018us$dmar)/length(nat2018us$dmar),3) # 61% sont mariés

##
## 1 2
## 0.611 0.267

round(table(nat2018us$priorlive)/length(nat2018us$priorlive),3) # 38% n'ont
pas eu d'enfant mais presque 30% ont déjà 2 enfants ou plus contre 12%

##
## 0 1 2 3 4 5 6 7 8 9 10 11
## 0.383 0.333 0.170 0.069 0.025 0.010 0.005 0.002 0.001 0.001 0.000 0.000
## 12 13 14 15 16 17 18
## 0.000 0.000 0.000 0.000 0.000 0.000 0.000

round(table(nat2018us$dplural)/length(nat2018us$dplural),3) # seul 3% des
mères ont donné naissance à des jumeaux

##
## 1 2 3 4 5
## 0.966 0.033 0.001 0.000 0.000
```

Santé de l'enfant

Score Apgar : 3.76, 50% au moins 9 1 A score of 0-3 2 A score of 4-6 3 A score of 7-8 4 A score of 9-10 5 Unknown or not stated

Poids à la naissance : 0227-8165 Number of grams 9999 Not stated birth weight

```
summary(select(pop,apgar5r))
```

```
##      apgar5r
## Min.      :1.000
## 1st Qu.:4.000
## Median :4.000
## Mean    :3.775
## 3rd Qu.:4.000
## Max.    :4.000
```

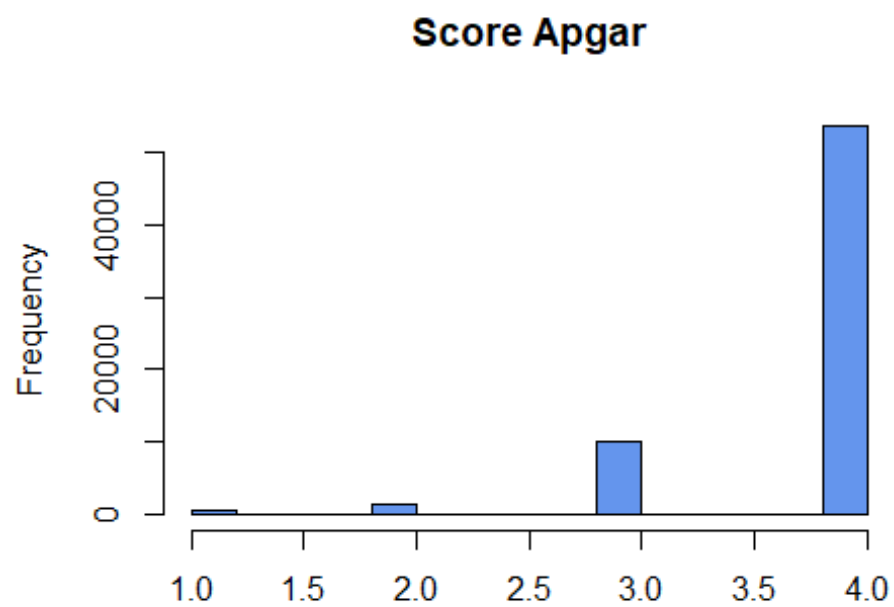
```
summary(select(pop,dbwt))
```

```
##      dbwt
## Min.    : 227
## 1st Qu.:2680
## Median :3172
## Mean    :3078
```

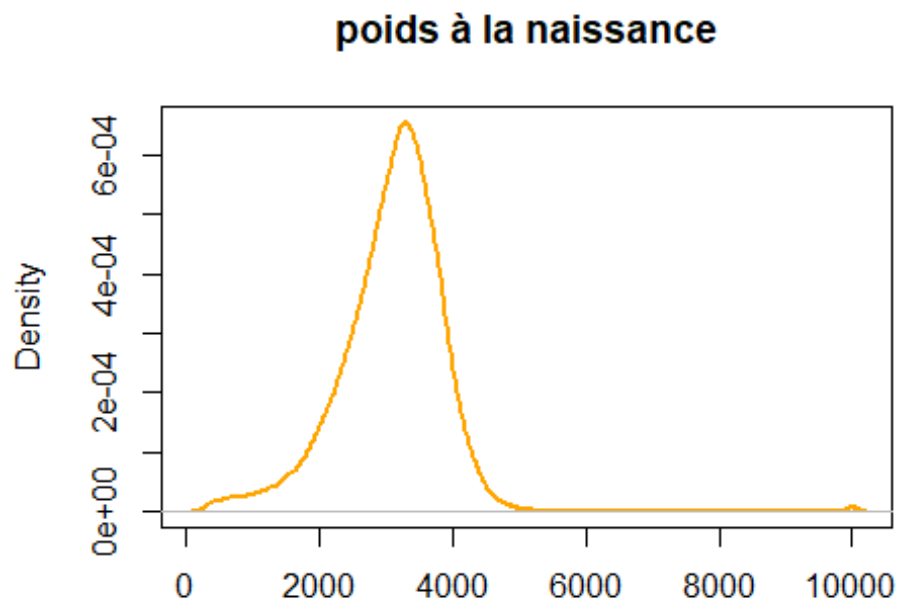
```
## 3rd Qu.:3565
```

```
## Max. :9999
```

```
hist(pop$apgar5r, col = "cornflowerblue", xlab = "", main = "Score Apgar")
```



```
plot(density(pop$dbwt), lwd = 2, col = "orange", xlab = "", main = "poids à  
la naissance")
```

Le poids à la naissance est plus élevé en moyenne mais le score Apgar est comparable

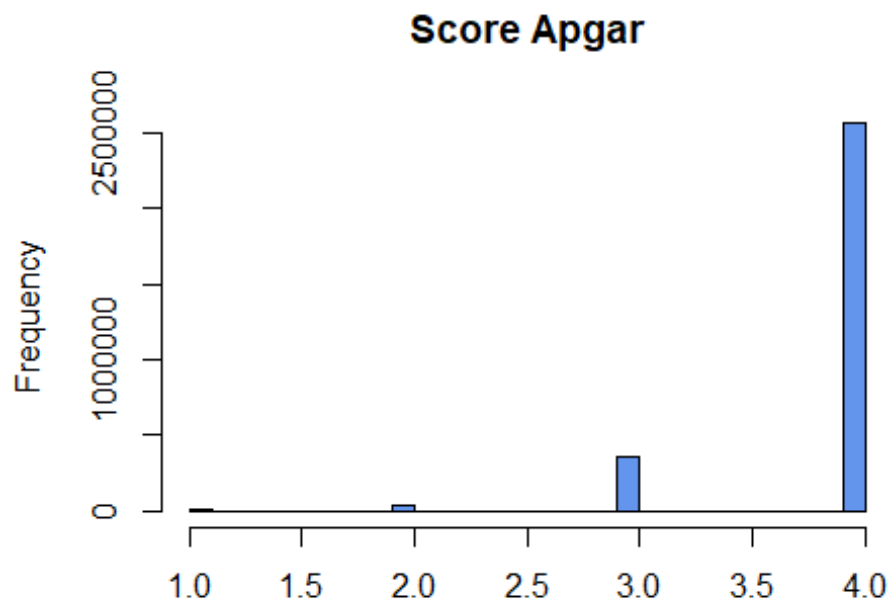
```
summary(select(nat2018us,apgar5r))
```

```
##      apgar5r
## Min.      :1.000
## 1st Qu.:4.000
## Median :4.000
## Mean      :3.839
## 3rd Qu.:4.000
## Max.      :4.000
```

```
summary(select(nat2018us,dbwt))
```

```
##      dbwt
## Min.      : 227
## 1st Qu.:2985
## Median :3320
## Mean      :3290
## 3rd Qu.:3650
## Max.      :9999
```

```
hist(nat2018us$apgar5r, col = "cornflowerblue", xlab = "", main = "Score  
Apgar")
```



```
plot(density(nat2018us$dbwt), lwd = 2, col = "orange", xlab = "", main =  
"poids à la naissance")
```

