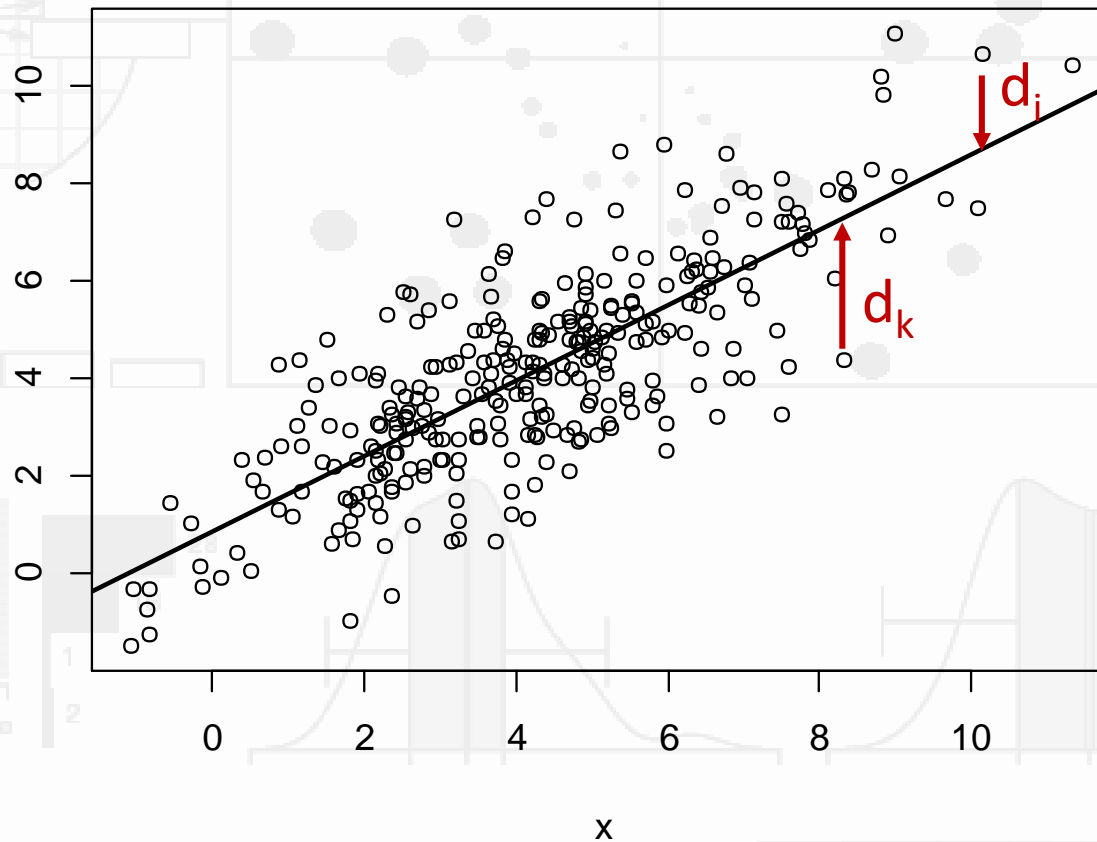


Régression linéaire « simple »

$d_1^2 + \dots + d_n^2$ est minimum



Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past: int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age: int 31 49 50 47 23 34 24 52 42 45 ...
 $ prof: Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree: int 4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip: int 0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant: int 2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratrerie: int 4 3 2 6 6 2 3 9 12 5 ...
 $ ecole: int 1 2 2 1 1 2 1 2 1 2 ...
 $ separation: int 0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant: int 0 0 0 0 NA 0 1 0 1 0 ...
 $ place: int 0 0 0 1 1 0 1 0 0 0 ...
 $ abus: int 0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons: int 1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons: int 0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons: int 1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons: int 0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons: int 0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons: int 0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons: int 0 0 0 0 0 0 0 0 0 0 ...
 $ char: int 1 1 1 1 1 1 1 1 4 1 ...
 $ rs: int 2 2 2 2 2 1 3 2 3 2 ...
 $ ed: int 1 2 3 2 2 2 3 2 3 2 ...
 $ dr: int 1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s: int 0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr: int 0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past: int 0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv: int NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char      : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs        : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed        : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr        : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratrerie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past: int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```


Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant  : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratrerie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratr  : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separa  : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char      : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs        : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed        : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr        : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame':  799 obs. of  26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past: int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```


Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame':  799 obs. of  26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/  8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratr  : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separa   : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enf : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons: int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons: int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons: int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s: int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr: int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past: int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv: int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```


Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int   4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int   0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant  : int   2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratr   : int   4 3 2 6 6 2 3 9 12 5 ...
 $ ecole     : int   1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int   0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int   0 0 0 0 NA 0 1 0 1 0 ...
 $ place     : int   0 0 0 1 1 0 1 0 0 0 ...
 $ abus      : int   0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int   1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int   0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int   1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int   0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int   0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int   0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int   0 0 0 0 0 0 0 0 0 0 ...
 $ char      : int   1 1 1 1 1 1 1 1 4 1 ...
 $ rs        : int   2 2 2 2 2 1 3 2 3 2 ...
 $ ed        : int   1 2 3 2 2 2 3 2 3 2 ...
 $ dr        : int   1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int   0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int   0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int   0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratie : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separation : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enfant : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons : int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons  : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons  : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons  : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons : int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s : int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr : int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past : int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv : int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Le fichier smp.l

Introduction à la statistique avec R > Régression linéaire



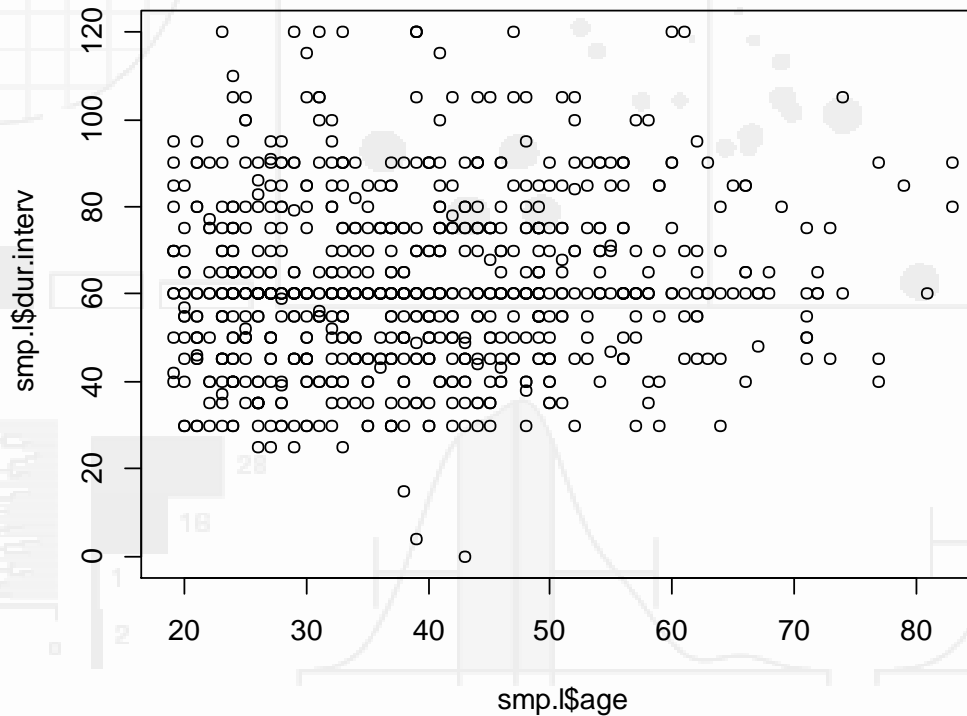
```
> smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
> str(smp.l)
'data.frame': 799 obs. of 26 variables:
 $ age      : int  31 49 50 47 23 34 24 52 42 45 ...
 $ prof     : Factor w/ 8 levels "agriculteur",...: 3 NA 7...
 $ duree    : int  4 NA 5 NA 4 NA NA 5 4 NA ...
 $ discip   : int  0 0 0 0 1 0 0 0 1 0 ...
 $ n.enfant : int  2 7 2 0 1 3 5 2 1 2 ...
 $ n.fratr  : int  4 3 2 6 6 2 3 9 12 5 ...
 $ ecole    : int  1 2 2 1 1 2 1 2 1 2 ...
 $ separa   : int  0 1 0 1 1 0 1 0 1 0 ...
 $ juge.enf : int  0 0 0 0 NA 0 1 0 1 0 ...
 $ place    : int  0 0 0 1 1 0 1 0 0 0 ...
 $ abus     : int  0 0 0 0 0 0 0 0 1 1 ...
 $ grav.cons: int  1 2 2 1 2 1 5 1 5 5 ...
 $ dep.cons : int  0 0 0 0 1 0 1 0 1 0 ...
 $ ago.cons : int  1 0 0 0 0 0 0 0 0 0 ...
 $ ptsd.cons: int  0 0 0 0 0 0 0 0 0 0 ...
 $ alc.cons : int  0 0 0 0 0 0 0 0 1 1 ...
 $ subst.cons: int  0 0 0 0 0 0 1 0 1 0 ...
 $ scz.cons : int  0 0 0 0 0 0 0 0 0 0 ...
 $ char     : int  1 1 1 1 1 1 1 1 4 1 ...
 $ rs       : int  2 2 2 2 2 1 3 2 3 2 ...
 $ ed       : int  1 2 3 2 2 2 3 2 3 2 ...
 $ dr       : int  1 1 2 2 2 1 2 2 1 2 ...
 $ suicide.s: int  0 0 0 1 0 0 3 0 4 0 ...
 $ suicide.hr: int  0 0 0 0 0 0 1 0 1 0 ...
 $ suicide.past: int  0 0 0 0 1 0 1 0 1 0 ...
 $ dur.interv: int  NA 70 NA 105 NA NA 105 84 78 60 ...
```

Régression linéaire simple

Introduction à la statistique avec R > Régression linéaire



```
> plot(smp.l$age, smp.l$dur.interv)
```

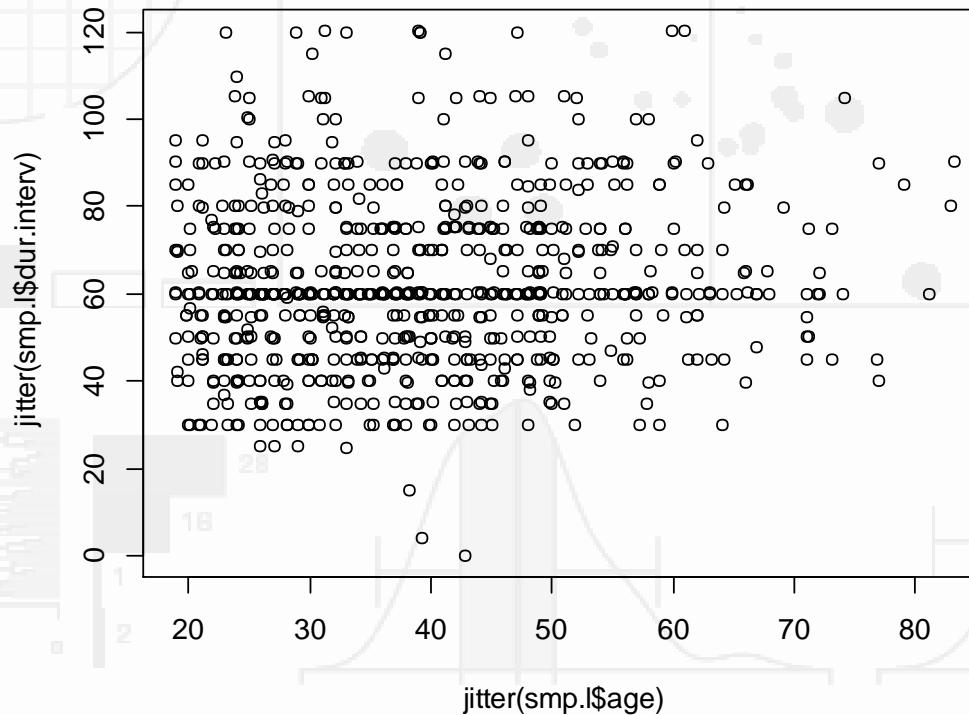


Régression linéaire simple

Introduction à la statistique avec R > Régression linéaire



```
> plot(jitter(smp.l$age),  
       jitter(smp.l$dur.interv))
```

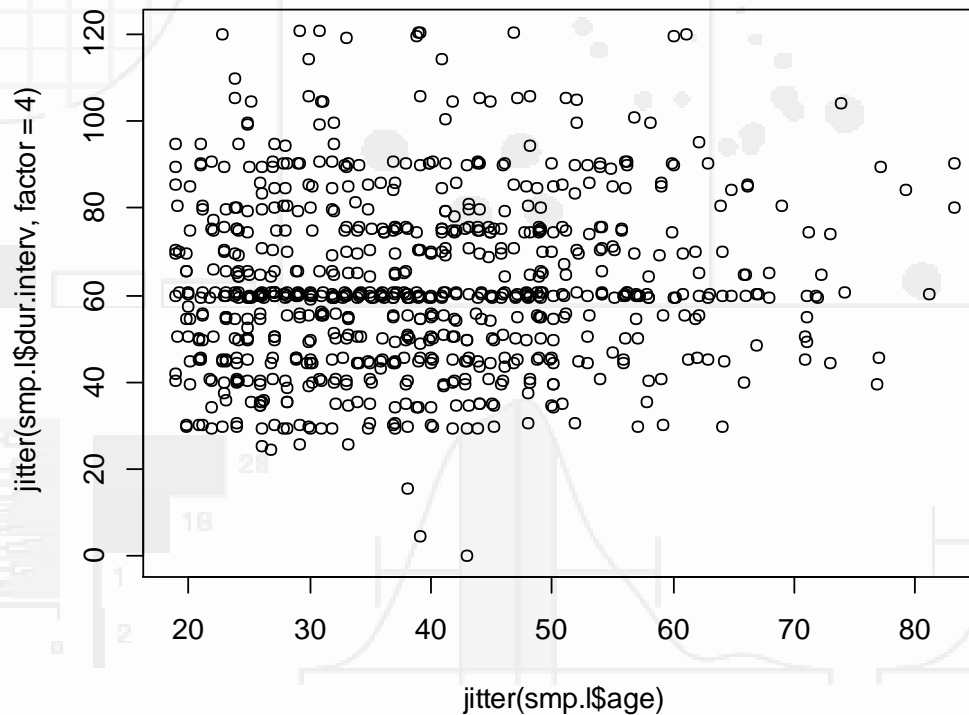


Régression linéaire simple

Introduction à la statistique avec R > Régression linéaire



```
> plot(jitter(smp.l$age),  
       jitter(smp.l$dur.interv, factor=4))
```

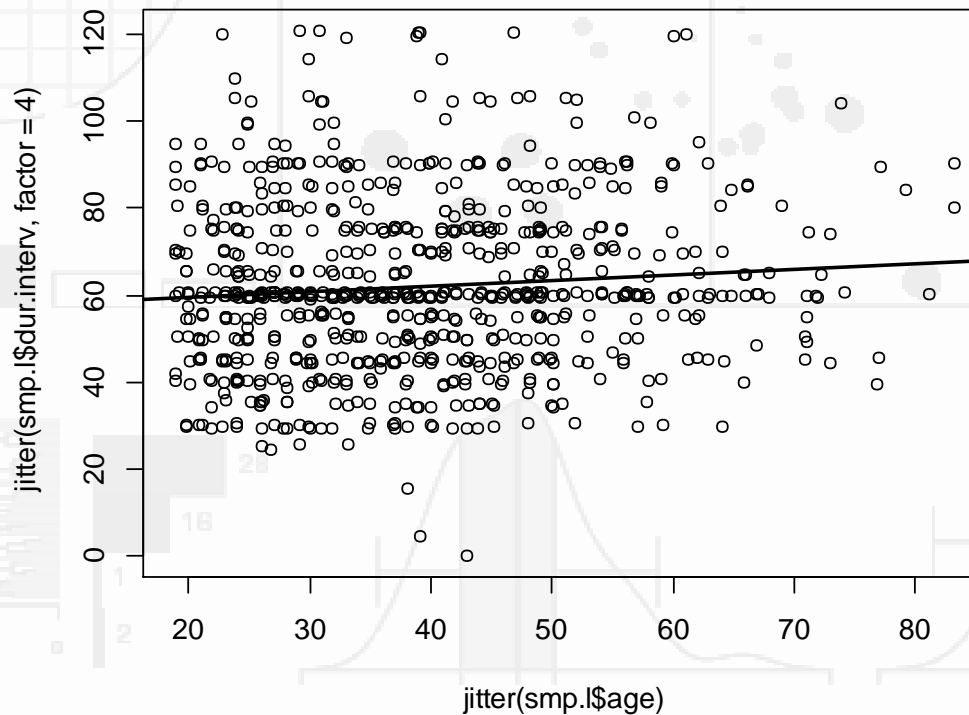


Régression linéaire simple

Introduction à la statistique avec R > Régression linéaire



```
> abline(lm(smp.l$dur.interv~smp.l$age),  
lwd=2)
```

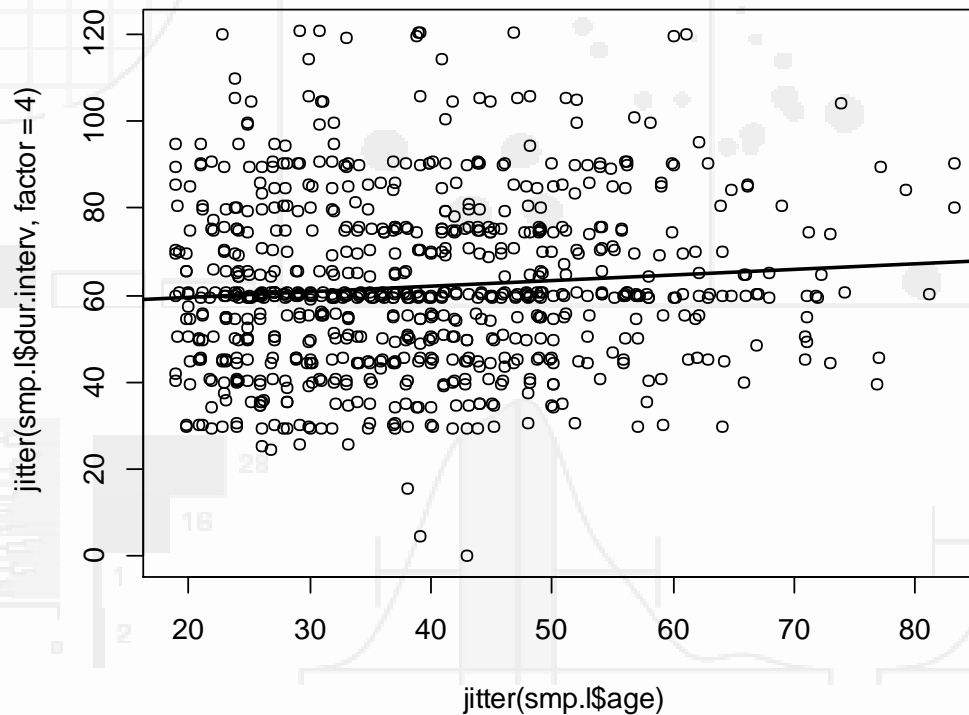


Régression linéaire simple

Introduction à la statistique avec R > Régression linéaire



$\text{durée} = a + b \times \text{age} \rightarrow b \neq 0 ?$



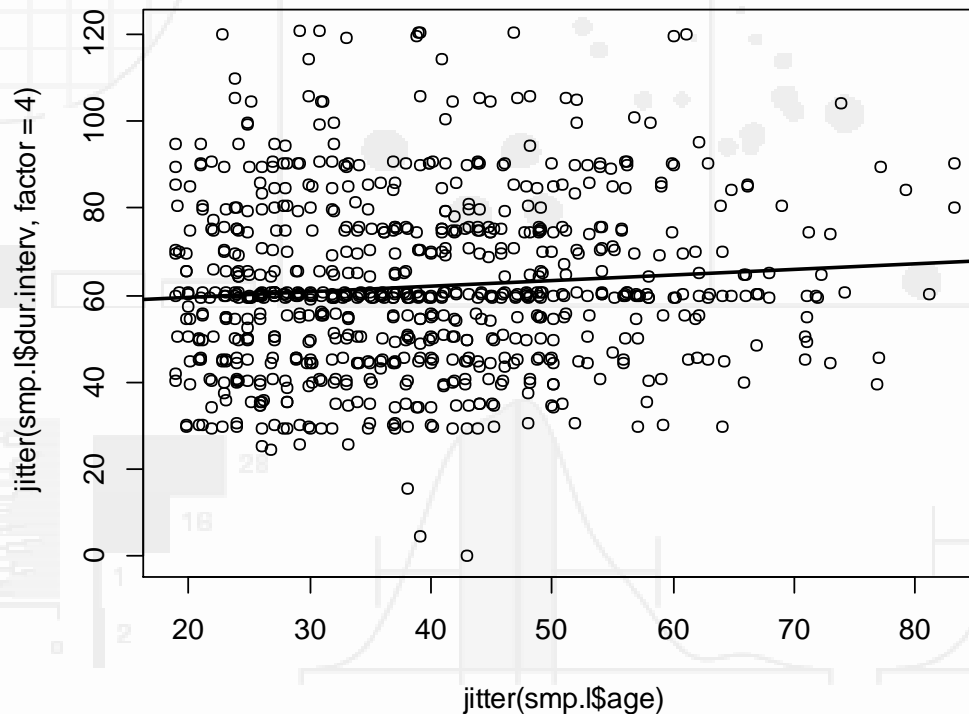
Régression linéaire simple

Introduction à la statistique avec R > Régression linéaire



$$\text{durée} = a + b \times \text{age} \rightarrow b \neq 0 ?$$

Modèle de régression : durée = a + b × age + bruit



$$\text{durée} = a + b \times \text{age} + \text{bruit}$$

```
> mod1 <- lm(dur.interv~age,data=smp.1)
> summary(mod1)
```

Call:

```
lm(formula = dur.interv ~ age, data = smp.1)
```

Residuals:

Min	1Q	Median	3Q	Max
-62.470	-14.402	-1.712	12.341	60.055

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	57.04091	2.22028	25.691	<2e-16 ***
age	0.12625	0.05375	2.349	0.0191 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 19.57 on 745 degrees of freedom

(52 observations deleted due to missingness)

Multiple R-squared: 0.00735, Adjusted R-squared: 0.006018

F-statistic: 5.516 on 1 and 745 DF, p-value: 0.0191

$$\text{durée} = a + b \times \text{age} + \text{bruit}$$

```
> mod1 <- lm(dur.interv~age,data=smp.1)
> summary(mod1)
```

Call:

```
lm(formula = dur.interv ~ age, data = smp.1)
```

Residuals:

Min	1Q	Median	3Q	Max
-62.470	-14.402	-1.712	12.341	60.055

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	57.04091	2.22028	25.691	<2e-16 ***
age	0.12625	0.05375	2.349	0.0191 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 19.57 on 745 degrees of freedom

(52 observations deleted due to missingness)

Multiple R-squared: 0.00735, Adjusted R-squared: 0.006018

F-statistic: 5.516 on 1 and 745 DF, p-value: 0.0191

$$\text{durée} = a + b \times \text{age} + \text{bruit}$$

```
> mod1 <- lm(dur.interv~age,data=smp.1)
> summary(mod1)
```

Call:

```
lm(formula = dur.interv ~ age, data = smp.1)
```

Residuals:

Min	1Q	Median	3Q	Max
-62.470	-14.402	-1.712	12.341	60.055

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	57.04091	2.22028	25.691	<2e-16 ***
age	0.12625	0.05375	2.349	0.0191 * b ≠ 0 ?

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 19.57 on 745 degrees of freedom

(52 observations deleted due to missingness)

Multiple R-squared: 0.00735, Adjusted R-squared: 0.006018

F-statistic: 5.516 on 1 and 745 DF, p-value: 0.0191

Conclusion

Introduction à la statistique avec R > Au-delà de la corrélation et du test t



```
smp.l <- read.csv2("D:/MOOC/Data/smp2.csv")
str(smp.l)

plot(smp.l$age, smp.l$dur.interv)
plot(jitter(smp.l$age), jitter(smp.l$dur.interv))
plot(jitter(smp.l$age), jitter(smp.l$dur.interv, factor=4))
abline(lm(smp.l$dur.interv~smp.l$age), lwd=2)

mod1 <- lm(dur.interv~age, data=smp.l)
summary(mod1)
```