Software de datos

```
  #Archivo de datos "malaria.csv"
  library(reshape2)
  malaria.raw <- read.csv("malaria.csv", header = FALSE, skip = 2) #Datos en br
uto
  head(malaria.raw)
```

```
##            V1      V2      V3      V4      V5      V6      V7      V8
## 1 Afghanistan   39263   54840   77549   69397   64880   81574   92202
## 2     Algeria      16      59       4      12       4       4      27
## 3      Angola 1999868 1496834 1632282 1682870 1573422 1377992 1533485
## 4   Argentina       0       0       0      26      86     130     387
## 5     Armenia      NA      NA      NA       0       0       0       0
## 6  Azerbaijan       0       3       4      50      78      72     109
##        V9    V10    V11    V12    V13  V14   V15
## 1   86129 116444 242022 360940 415356   NA 94475
## 2       1      2      3      6      8    8    35
## 3 1082398 889572     NA     NA     NA   NA    NA
## 4     212    252    115    122    125  215   440
## 5       0      7     47     29     52   79   141
## 6     143    242    386    482    506 1058  1526
```

```
  cnames<-read.csv("malaria.csv", header = FALSE, skip = 1, nrows=1)

  cnames <- sapply(cnames, function(x){x <- trimws(as.character(x))})
  names(malaria.raw) <- cnames

  # Solo filas sin valores NA
  nrow(malaria.raw[!complete.cases(malaria.raw),])
```

```
## [1] 47
```

```
  malaria <- malaria.raw[complete.cases(malaria.raw),]

  # Convierto las columnas de los años en filas
  malaria<-melt(malaria.raw)
```

```
## Using Country as id variables
```

```
  head(malaria)
```

```
##       Country variable   value
## 1 Afghanistan     2013   39263
## 2     Algeria     2013      16
## 3      Angola     2013 1999868
## 4   Argentina     2013       0
## 5     Armenia     2013      NA
## 6  Azerbaijan     2013       0
```

```
  names(malaria) <- c("country","year","cases")

  malaria$country <- as.character(malaria$country)
  malaria$year <- as.character(malaria$year)
  head(malaria)
```

```
##       country year   cases
## 1 Afghanistan 2013   39263
## 2     Algeria 2013      16
## 3      Angola 2013 1999868
## 4   Argentina 2013       0
## 5     Armenia 2013      NA
## 6  Azerbaijan 2013       0
```

```
  #Archivo de datos "water.csv"
  water.raw <- read.csv("water.csv", header = FALSE,skip = 2)
  head(water.raw)
```

```
##            V1   V2 V3 V4 V5 V6 V7 V8
## 1 Afghanistan 2012 56 90 64 23 47 29
## 2 Afghanistan 2010 50 81 57 23 44 28
## 3 Afghanistan 2005 34 59 40 22 38 26
## 4 Afghanistan 2000 18 36 22 21 32 23
## 5 Afghanistan 1995  3 14  5 20 26 21
## 6 Afghanistan 1990  3 NA NA NA NA NA
```

```
  cnames <- read.csv("water.csv", header = FALSE, skip = 1, nrows=1)
  cnames <- sapply(cnames, function(x){x<-trimws(as.character(x))})

  names(water.raw) <- cnames
  names(water.raw) <- tolower(names(water.raw))

  water.raw$country <- as.character(water.raw$country)

  names(water.raw)[3:5] <-paste(names(water.raw)[4:6],sep = "",".drink")
  names(water.raw)[6:8] <- c("rural.facilities","urban.facilities","total.facil
ities")

  # Solo filas sin valores NA
  nrow(water.raw[!complete.cases(water.raw),])
```

```
## [1] 95
```

```
  water <- water.raw[complete.cases(water.raw),]

  head(water)
```

```
##       country year urban.drink total.drink rural.drink rural.facilities
## 1 Afghanistan 2012          56          90          64                23
## 2 Afghanistan 2010          50          81          57                23
## 3 Afghanistan 2005          34          59          40                22
## 4 Afghanistan 2000          18          36          22                21
## 5 Afghanistan 1995           3          14           5                20
## 7     Albania 2012          94          97          96                86
##   urban.facilities total.facilities
## 1               47               29
## 2               44               28
## 3               38               26
## 4               32               23
## 5               26               21
## 7               95               91
```

```
  #Se fusionan los datos por los campos comunes "country" y "year"
  malaria.water <- merge(malaria,water,by=c("country","year"))

  head(malaria.water)
```

```
##       country year  cases urban.drink total.drink rural.drink
## 1 Afghanistan 2000  94475          18          36          22
## 2 Afghanistan 2005 116444          34          59          40
## 3 Afghanistan 2010  69397          50          81          57
## 4 Afghanistan 2012  54840          56          90          64
## 5     Algeria 2000     35          84          93          89
## 6     Algeria 2005      2          81          88          86
##   rural.facilities urban.facilities total.facilities
## 1               21               32               23
## 2               22               38               26
## 3               23               44               28
## 4               23               47               29
## 5               82               99               92
## 6               86               98               94
```