Data Import

Quantide Srl

2016-05-02

First of all, set your working directory in the data folder, using `setwd()` function, like in this example

```r
setwd("C:/Users/Veronica/Documents/data)
```

Text Files

Exercise 1

a. Import text file named "tuscany.txt" and save it in the object `tuscany_df`.
   Open the text file before importing it to control if the first row contains column names and to control the field and the decimal separator characters. Remember to not import the character columns as factors.

```r
tuscany_df <- read.table("tuscany.txt", header = TRUE, sep = "|",
                         dec=".", stringsAsFactors = FALSE, quote = "")
tuscany_df
```

```
##    id sex year_of_birth marital_status  income house_number
## 1   1   M          1969        married 16101.1       5144.0
## 2   2   M          1962         single 17220.0       6158.0
## 3   3   M          1965        divorcee 28801.9      10078.0
## 4   4   F          1968         single 25964.0      11133.7
## 5   5   M          1975        married 16522.5       5078.0
## 6   6   M          1977        married 18124.0       5115.0
## 7   7   M          1972         single 25145.0       9801.0
## 8   8   F          1974        married 12741.0       5752.0
## 9   9   F          1978        married 19051.0       6947.0
## 10 10   F          1964        married 28893.0       9635.0
## 11 11   M          1963        divorcee 29389.0       9950.0
## 12 12   F          1975        married 20169.0       7248.0
## 13 13   F          1984        married 20874.0       5313.0
## 14 14   M          1971        married 25649.4       8391.0
## 15 15   F          1969         single 27457.0       9396.0
## 16 16   F          1977        married 19975.0       6717.0
## 17 17   F          1968       marriedo 20824.0       5699.0
## 18 18   F          1956        married 13890.0       4479.0
## 19 19   F          1972        married 25567.0       9258.0
## 20 20   F          1973         single 13551.0       3890.0
## 21 21   F          1972        divorcee 15954.0       6142.0
## 22 22   M          1961        divorcee 27277.0       8470.0
## 23 23   M          1970        married 12444.0       3840.0
## 24 24   M          1979        married 22243.0       6473.0
## 25 25   F          1974        married 25606.0       6583.0
## 26 26   M          1978        divorcee 23225.7       6261.0
## 27 27   F          1975        married 16617.0       5080.0
```

```
## 28 28   F          1951       married 30496.0      10760.0
## 29 29   M          1964       married 26913.0       8785.0
## 30 30   F          1963        single 27856.0       7801.0
##                 city_name province provincial_acronym
## 1              Riparbella     Pisa                 PI
## 2               Capolona   Arezzo                 AR
## 3              Pomarance     Pisa                 PI
## 4                 Cascina     Pisa                 PI
## 5                Quarrata  Pistoia                 PT
## 6  Castiglion Fiorentino   Arezzo                 AR
## 7    Castel San Niccol\xf2   Arezzo                 AR
## 8    Barberino Val d Elsa  Firenze                 FI
## 9          Forte dei Marmi    Lucca                 LU
## 10                Fauglia     Pisa                 PI
## 11       Capraia e Limite  Firenze                 FI
## 12        Gambassi Terme  Firenze                 FI
## 13                 Bucine   Arezzo                 AR
## 14            Terricciola     Pisa                 PI
## 15    Campiglia Marittima  Livorno                 LI
## 16                Abetone  Pistoia                 PT
## 17                 Vaglia  Firenze                 FI
## 18         Bagno a Ripoli  Firenze                 FI
## 19              Pontedera     Pisa                 PI
## 20         Rio nell Elba  Livorno                 LI
## 21                  Signa  Firenze                 FI
## 22               Stazzema    Lucca                 LU
## 23     San Piero a Sieve  Firenze                 FI
## 24               Piombino  Livorno                 LI
## 25              Molazzana    Lucca                 LU
## 26             Santa Luce     Pisa                 PI
## 27       Castelfiorentino  Firenze                 FI
## 28                Livorno  Livorno                 LI
## 29              Gallicano    Lucca                 LU
## 30                Fiesole  Firenze                 FI
```

b. Visualize the first rows of `tuscany_df`

```
head(tuscany_df)
```

```
##   id sex year_of_birth marital_status  income house_number
## 1  1   M          1969        married 16101.1       5144.0
## 2  2   M          1962         single 17220.0       6158.0
## 3  3   M          1965        divorcee 28801.9      10078.0
## 4  4   F          1968         single 25964.0      11133.7
## 5  5   M          1975        married 16522.5       5078.0
## 6  6   M          1977        married 18124.0       5115.0
##                 city_name province provincial_acronym
## 1              Riparbella     Pisa                 PI
## 2               Capolona   Arezzo                 AR
## 3              Pomarance     Pisa                 PI
## 4                 Cascina     Pisa                 PI
## 5                Quarrata  Pistoia                 PT
## 6  Castiglion Fiorentino   Arezzo                 AR
```

Exercise 2

Import 7 rows of the text file named "solar.txt" skipping the first two rows and save it in the object `solar_df`. Open the text file before importing it to control if the first row contains column names and to control the field and the decimal separator characters. Remember to not import the character columns as factors.

```r
solar_df <- read.table("solar.txt", header = FALSE, sep = ",",
                       dec=".", stringsAsFactors = FALSE, quote = "",
                       nrows = 7, skip = 2)
solar_df
```

```
##     V1    V2    V3    V4
## 1 mar 23877 24671 22455
## 2 apr 24377 23677 23670
## 3 mag 24581 25476 24999
## 4 giu 22154 21998 22451
## 5 lug 20924 21645 23871
## 6 ago 23183 22576 23556
## 7 set 27446 27695 28664
```

Exercise 3

Save the following data frame in a .txt file named "exercise-3.txt" in data folder.

```r
df <- data.frame(col1=1:4, col2=4:1, col3=c("one", "two", "three", "four"),
                 stringsAsFactors = FALSE)
```

```r
write.table(df, file="exercise-3.txt")
```