

# Introduzione al software R

*Metodi Quantitativi per Economia,  
Finanza e Management*

*Esercitazione n°1*

# Orario di ricevimento

<b>Alberto Saccardi</b>	<a href="mailto:alberto.saccardi@nunatac.it">alberto.saccardi@nunatac.it</a> <a href="mailto:asaccardi@liuc.it">asaccardi@liuc.it</a>	Lunedì 17-18 <i>Aula lezione</i>
<b>Anna Pozzebon</b>	<a href="mailto:apozzebon@liuc.it">apozzebon@liuc.it</a>	Venerdì 10.30-11.00 e 12.30 – 13.00 <i>Laboratorio Grande</i>
<b>Davide Giannuzzi</b>	<a href="mailto:dgiannuzzi@liuc.it">dgiannuzzi@liuc.it</a>	Venerdì 10.30-11.00 e 12.30 – 13.00 <i>Laboratorio Grande</i>

# Suddivisione per Esercitazioni

**Venerdì ore 09.00 - 10.30**

Studenti di Banche, Amministrazione e Marketing.

**Venerdì ore 11.00 - 12.30**

Studenti di EDI e Risorse umane.

# Metodi Quantitativi per Economia, Finanza e Management

## Modalità d'esame:

- ❖ **Frequentanti** (è necessario raggiungere il 70% di presenze tra lezioni/esercitazioni, **obbligo di firma del foglio presenze**)
  - ✓ Lavoro di Gruppo – *punteggio massimo 21 punti, la sufficienza è raggiunta se si ottiene un punteggio  $\geq 13$ .*
  - ✓ Test a risposte multiple (15 min) – *10 domande con 5 modalità di risposta. Punteggio massimo 10 punti, la sufficienza è raggiunta se si ottiene un punteggio  $\geq 5$ . Il test può essere ripetuto un massimo di 2 volte (nel caso di ritiro durante l'esame non viene conteggiato)*

# Metodi Quantitativi per Economia, Finanza e Management

## ❖ Non frequentanti:

- Chi non ha raggiunto il 70% delle firme presenza
- Chi ha svolto il lavoro di gruppo con un punteggio  $<13/21$
- Chi ha svolto 2 volte il test a risposte multiple  $<5/10$

## Modalità d'esame:

- ✓ Test a risposte multiple (15 min) - 10 domande con 5 modalità di risposta. Punteggio massimo 10 punti, la sufficienza è raggiunta se si ottiene un punteggio  $\geq 5$ .
- ✓ Test a domande aperte (45 min) – punteggio massimo 21 punti, la sufficienza è raggiunta se si ottiene un punteggio  $\geq 13$ .

# Metodi Quantitativi per Economia, Finanza e Management

## Lavoro di Gruppo (min 3 – max 4 persone):

- ✓ prova di natura applicativa prevista **solo** per i frequentanti;
- ✓ consiste nell'elaborazione e nell'analisi quantitativa (mediante l'utilizzo di R) di un set di dati raccolti mediante una **survey** opportunamente predisposta dal gruppo di lavoro;
- ✓ qualora la prova abbia esito sufficiente, avrà validità di un anno accademico. Sono previsti al massimo 2 tentativi per la modalità frequentanti.
- ✓ la scelta della modalità «non frequentante» è irreversibile: non potranno più essere utilizzati eventuali tentativi residui da frequentante.

Fase raccolta dati:  
Questionario di  
rilevazione

Costruzione base  
dati

Analisi Dati

Report cartaceo  
(formato pptx)

# Metodi Quantitativi per Economia, Finanza e Management

## Lavoro di Gruppo (min 3 – max 4 persone):

- La composizione del gruppo dovrà essere comunicata tramite posta elettronica, entro il **05/10/2017**, a [apozzebon@liuc.it](mailto:apozzebon@liuc.it) e [dgiannuzzi@liuc.it](mailto:dgiannuzzi@liuc.it).
- Inviare una email con le seguenti informazioni:
  - nome, cognome e numero di matricola dei partecipanti (massimo 4 componenti per gruppo)
  - nome del gruppo
  - titolo/argomento del lavoro di gruppo che si intende presentare

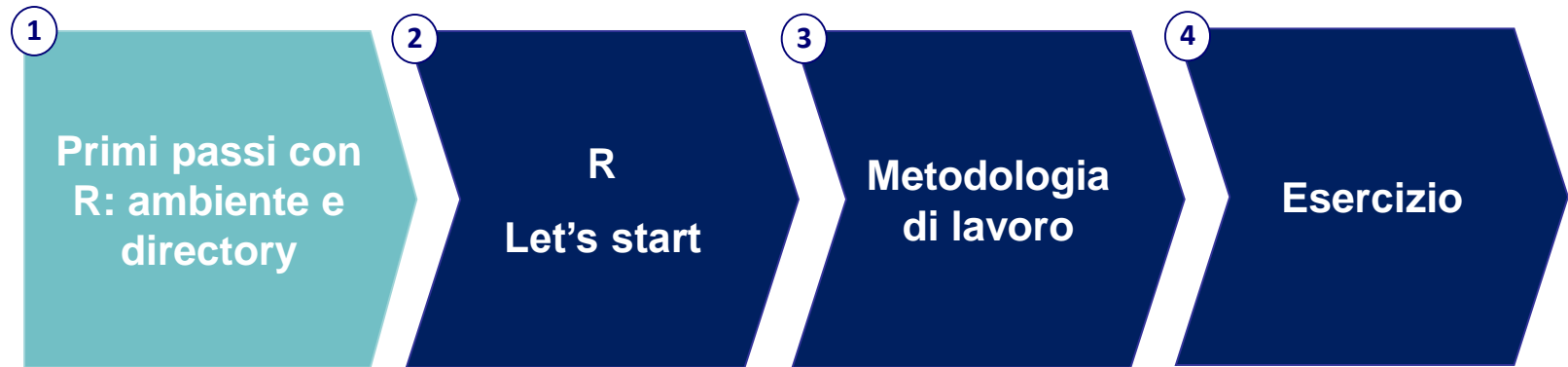
# Metodi Quantitativi per Economia, Finanza e Management

Tipo di analisi	Strumenti	Lavoro di gruppo
ANALISI UNIVARIATA	<ul style="list-style-type: none"> <li>- DISTRIBUZIONI DI FREQUENZA</li> <li>- INDICI DI POSIZIONE (MISURE DI TENDENZA CENTRALE E MISURE DI TENDENZA NON CENTRALE)</li> <li>- INDICI DI DISPERSIONE</li> <li>- MISURE DI FORMA DELLA DISTRIBUZIONE</li> </ul>	<b>OBBLIGATORIO</b>
ANALISI BIVARIATA E TEST STATISTICI PER LO STUDIO DELL'ASSOCIAZIONE TRA VARIABILI	<p><b><u>Due variabili qualitative o quantitative discrete:</u></b>            TABELLA DI CONTINGENZA E INDICI CHI QUADRO E V DI CRAMER            TEST CHI QUADRO PER L'INDIPENDENZA STATISTICA</p> <p><b><u>Due variabili quantitative continue:</u></b>            INDICE DI CORRELAZIONE DI PEARSON (<math>\rho</math>) E COVARIANZA            TEST t PER L'INDIPENDENZA LINEARE</p> <p><b><u>Una variabile qualitativa e una quantitativa continua:</u></b>            INDICE <math>\eta^2</math>            TEST F PER L'INDIPENDENZA IN MEDIA</p>	<b>OBBLIGATORIO</b>
ANALISI MULTIVARIATA	<ul style="list-style-type: none"> <li>- ANALISI FATTORIALE</li> <li>- REGRESSIONE LINEARE</li> <li>- REGRESSIONE LOGISTICA</li> </ul>	<b>OBBLIGATORIO</b>



# Metodi Quantitativi per Economia, Finanza e Management

**Obiettivi di questa esercitazione:**



# R – Caratteristiche

R è un software statistico *Open Source* che consente di:

- ✓ Gestire e modificare dati/informazioni sotto forma tabellare
- ✓ Modificare o produrre nuovi dati tramite trasformazioni/funzioni
- ✓ Analizzare i dati con funzioni dedicate
- ✓ Produrre tabulati di tipo standard o personalizzato per la presentazione dei risultati



# R – Filosofia

Analisi statistica = output specifici, si analizza solo ciò che serve. L'analisi statistica è svolta attraverso una serie di passi, con risultati intermedi salvati in oggetti



Ogni oggetto può essere modificato, visualizzato, richiamato dall'utente ad ogni passo dell'analisi.

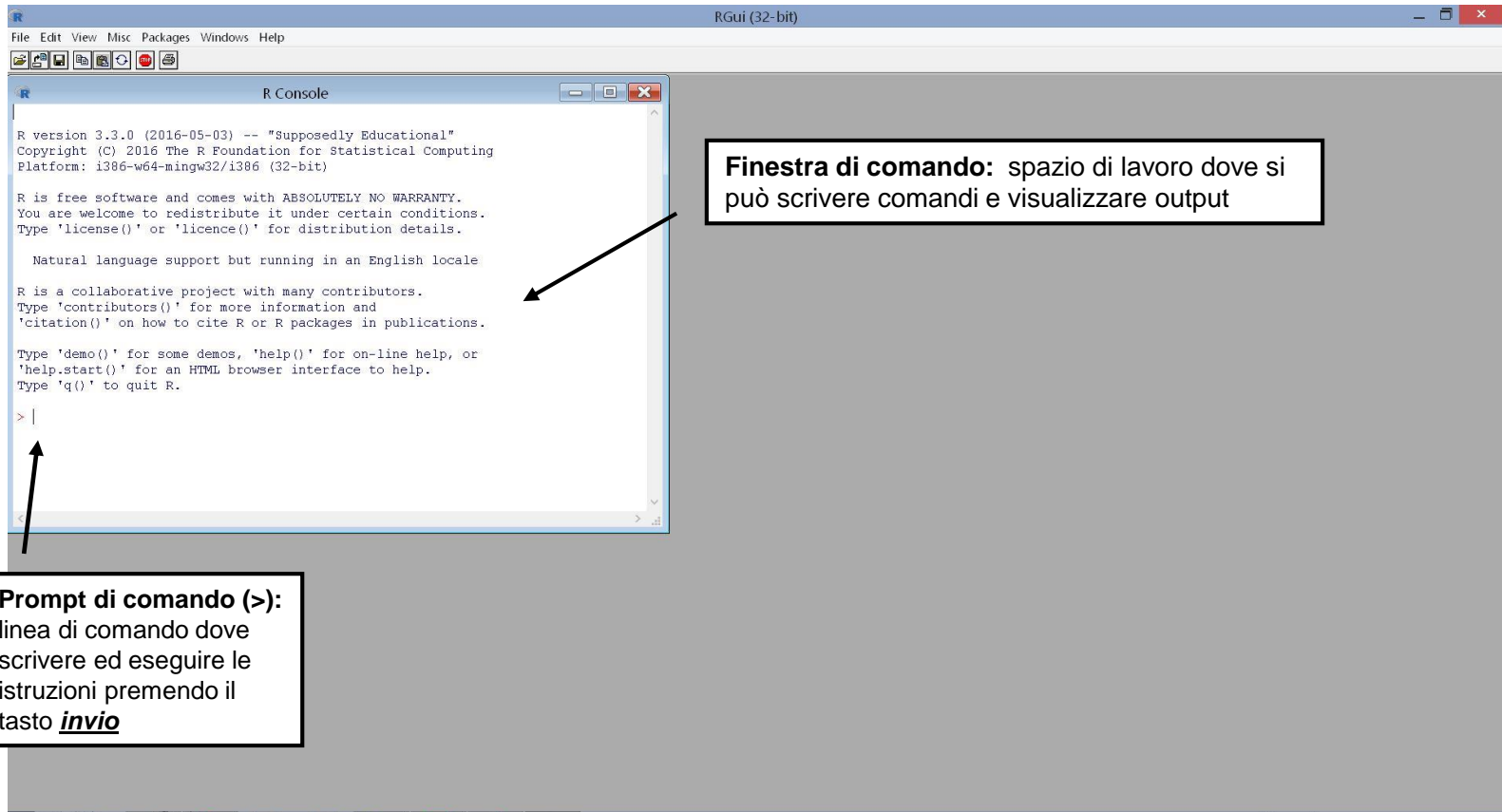


# Prima di iniziare...

- Controllare se sul pc su cui state lavorando esiste già una cartella C:\corso. In tal caso eliminare tutto il contenuto. In caso contrario creare la cartella **corso** all'interno del disco C
- Andare sul disco condiviso F nel percorso ***F:\corsi\Metodi\_Quantitativi\_EFM\_1819\esercitazione1*** e copiare il contenuto nella cartella C:\corso
- Aprire il programma R (Start → All Programs → R)

# R – Interfaccia

- Apertura della sessione: cliccare sull'icona R di Windows



# R – Interfaccia

1

The screenshot shows the R GUI (32-bit) window. The 'File' menu is open, highlighting 'New script...'. A callout box points to this menu item with the text: 'Click su File – New Script per aprire la finestra di R Editor'. Another callout box points to the continuation symbols (red '+') in the console window, stating: 'Continuation symbol (+): indica che la linea di codice appartiene ancora alla precedente linea di codice, ovvero che la funzione lanciata non è stata chiusa'. A third callout box points to the prompt symbol (>) and the cursor in the console, explaining: 'Prompt symbol: indica dove scrivere i comandi. E' caratterizzato dal simbolo > e dal cursore lampeggiante alla destra del prompt.'. A fourth callout box points to the 'Untitled - R Editor' window, describing it as: 'R Editor (Script): spazio di lavoro che raccoglie i comandi e le funzioni da eseguire (con F5) in modo tale da poterle salvare e recuperare il lavoro svolto una volta chiuso R'. A large red number '2' is placed in the background of the R Editor window. The Windows taskbar at the bottom shows various application icons, including Excel, Outlook, Chrome, SAS, and the R logo.

2

**Prompt symbol:** indica dove scrivere i comandi. E' caratterizzato dal simbolo > e dal cursore lampeggiante alla destra del prompt.

**Click su File – New Script** per aprire la finestra di R Editor

**Continuation symbol (+):** indica che la linea di codice appartiene ancora alla precedente linea di codice, ovvero che la funzione lanciata non è stata chiusa

**R Editor (Script):** spazio di lavoro che raccoglie i comandi e le funzioni da eseguire (con F5) in modo tale da poterle salvare e recuperare il lavoro svolto una volta chiuso R



# R – Directory (1/3)

## **Cos'è una directory R:**

Spazio sul disco fisso in cui sono salvati i file con i dati da analizzare (corrisponde alla cartella fisica di lavoro).

All'apertura di R la directory viene assegnata automaticamente

Per sapere quale sia la directory su cui R sta lavorando, digitiamo:

```
>getwd()
```

E l'output risultante, per esempio, è:

```
>getwd()
```

```
>[1] "C:/Programmi/R/Corso"
```



# R – Directory (2/3)

Per cambiare la directory di lavoro in:

*C:\Corso*

digitiamo il comando:

```
> setwd("C:/Corso")
```

**Osservazione:** mentre in Windows generalmente il simbolo che separa le cartelle successive entro un percorso è “\”, in R è sostituito da “/”.

Con il comando

```
> dir()
```

possiamo vedere tutti i file contenuti nella directory di lavoro.



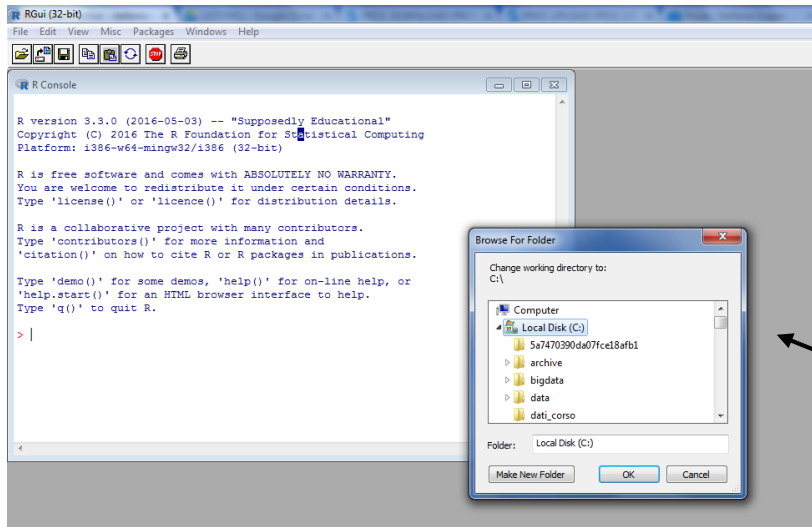
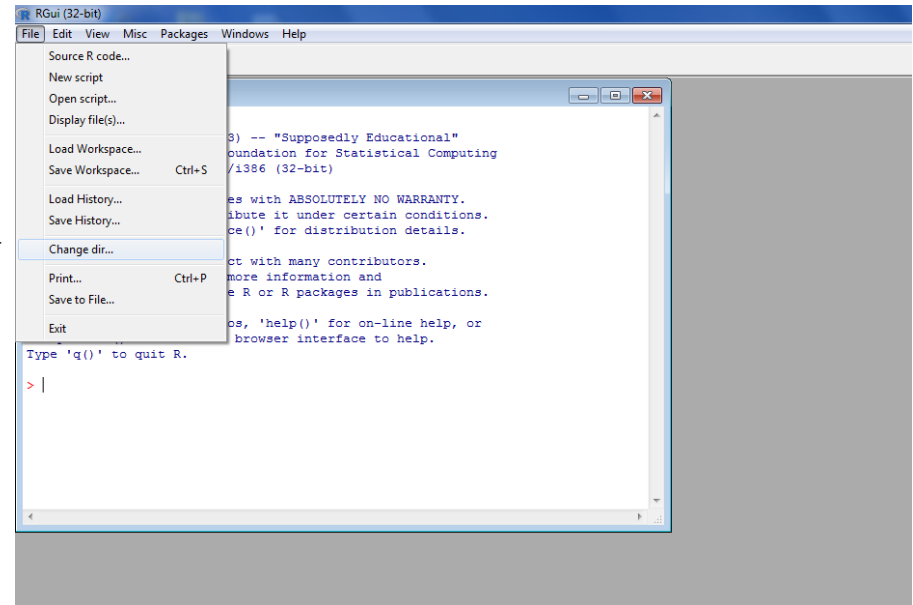


# R – Directory (3/3)

Un altro modo per cambiare la directory di lavoro è:

Click su File – Change Directory per aprire la finestra di ricerca path nuova directory

1



2

Si aprirà il **Browse for Folder**, dove si cercherà la cartella che contiene tutti i dataset che si vogliono importare – Click su **OK**

# Metodi Quantitativi per Economia, Finanza e Management

**Obiettivi di questa esercitazione:**



# R – Programmi (1/2)

Le istruzioni della sintassi R:

- Sono *case sensitive* (differenza tra maiuscole e minuscole)
- Possono estendersi su più linee del Program Editor
- Due comandi differenti devono stare su due righe differenti nel Prompt dei comandi
- I commenti nel Program Editor si aprono #



# R – Programmi (2/2)

L'esecuzione di un programma R avviene in tre step

1. creazione in R/importazione (ad es. da Excel) di una tabella contenente i dati
2. esecuzione di un codice R (ad es. una funzione)
3. produzione di un output (ad es. statistiche di sintesi sui dati) come risultato del codice

**N.B. I programmi R possono essere salvati durante qualsiasi momento della sessione di lavoro, per poi essere richiamati, sottomessi o modificati in sessioni successive.**



# R – Le variabili (1/3)

Regole per i nomi delle variabili in R:

1. Possono essere alfanumeriche (contenere sia numeri che lettere)
2. Possono contenere i simboli `_` (*underscore*) e `.` (*punto*)
3. Devono iniziare sempre con una lettera alfabetica o con il simbolo `.` (*punto*). Se il nome della variabile inizia con `.` (*punto*) il secondo carattere non può essere un numero
4. Non possono contenere spazi
5. Non possono essere parole speciali

Reserved words in R

if	else	repeat	while	function
for	in	next	break	TRUE
FALSE	NULL	Inf	NaN	NA
NA_integer_	NA_real_	NA_complex_ NA_character_.		



# R – Le variabili (2/3)

## Tipologia:

- Alfanumeriche: sulle quali è possibile compiere operazioni di confronto, ordinamento, concatenazione, selezione.
- Numeriche: sulle quali è possibile effettuare anche operazioni algebriche.

```
> x[,c("marca", "peso", "cilindrata")]
```

	marca	peso	cilindrata
1	alfa-romeo	1155.8	2.1303
2	alfa-romeo	1155.8	2.1303
3	alfa-romeo	1280.5	2.4908
4	audi	1060.1	1.7862
5	audi	1281.0	2.2286
6	audi	1137.2	2.2286

Comando di R

Output di R

Intestazione (prima riga dell'output), contiene i **nomi delle variabili** del dataset



# R – Le variabili (3/3)

## Valori mancanti (missing) :

Nel dataset di origine:

- se la variabile è di *tipo numerico*, R interpreta lo spazio vuoto del file di origine imputando il valore missing a NA (Not Available)
- Se la variabile è di tipo *character* bisogna inserire manualmente NA nel file d'origine. Nella tabella R il missing viene rappresentato con <NA>

**NB:** NA non può essere un attributo di una variabile categoriale (vedi *Tabella parole speciali R*)



# R – Le variabili (3/3)

## Valori mancanti (missing) :

	A	B	C	D	E	F	G
1	n_questionario	motivo_utilizzo_2	num_sms_e				
2	238	Piacere/Tempo libero	15				
3	6	Piacere/Tempo libero					
4	14	Piacere/Tempo libero	20				
5	44	NA	30				
6	68	Partner	70				
7	74	Famigliari	10				
8	75	Piacere/Tempo libero					
9	87	Partner	55				

Valore missing di una  
variabile Numerica (Excel)

Valore missing di una  
variabile Numerica (R)

Valore missing di una variabile  
Alfanumerica (Excel)

Valore missing di una variabile  
Alfanumerica (R)

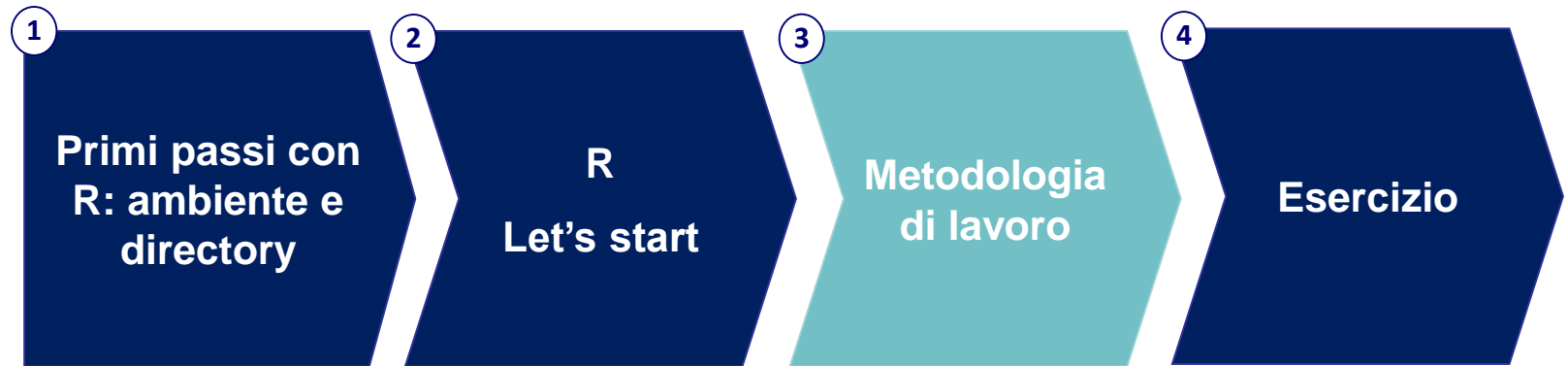
```
> x
  n_questionario  motivo_utilizzo_2 num_sms_e
1          238   Piacere/Tempo libero      15
2           6   Piacere/Tempo libero      NA
3          14   Piacere/Tempo libero      20
4          44   Piacere/Tempo libero      <NA>
5          68   Partner                 70
6          74   Famigliari                10
7          75   Piacere/Tempo libero      NA
8          87   Partner                    55
> |
```





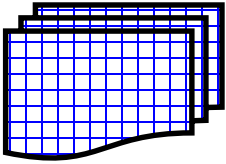
# Metodi Quantitativi per Economia, Finanza e Management

**Obiettivi di questa esercitazione:**



# Anteprima

Analisi



**DATI GREZZI**  
(\* .xls, \* .txt, ...)

**IMPORTAZIONE**

Tabella in R,  
Chiamato Data set o  
Data Frame

```
getwd ()
ESERCIZIO 2
isole<- islands
data.frame (isole)
hist (isole, main="istogramma")
mtext ("istogramma per le dimensioni delle isole")

barplot (isole)
fix (isole2)
isole2<-as.data.frame (isole)
isole2$names

isole2$names

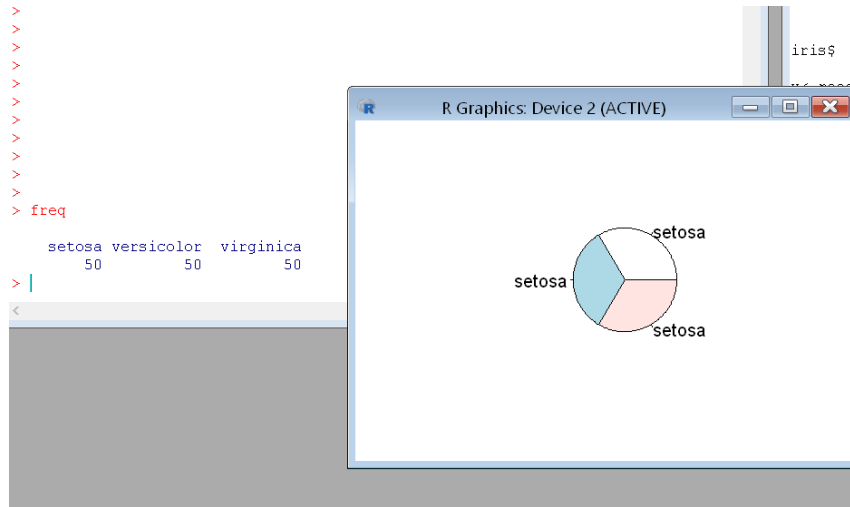
is.data.frame (isole2)
par (mfrow=c (1,2))

hist (isole, breaks="Scott")
hist (isole, breaks="sturges")

boxplot (split (isole2$row.names, isole2$isole))
boxplot (split (isole2$names, isole2$isole))
```

**ESECUZIONE PROGRAMMA**

**RISULTATI**



# Importazione file esterno (1/3)

È possibile procedere all'importazione di un file dati. Di seguito verrà illustrato il procedimento:

1. Salvare il file da importare in formato CSV
2. Due possibili comandi per importare un file CSV, che conducono allo stesso risultato :

```
read.csv('nome_file.csv', header=TRUE)
```

```
read.table('nome_file.csv', header=TRUE, sep=',')
```

Inserire, tra virgolette, il nome del file da importare (RICORDARE CHE R è CASE SENSITIVE). Nel caso in cui il file si trovi nella directory di lavoro non serve specificare il percorso fisico, altrimenti è necessario riportarlo, ricordando che in R la barra obliqua \ deve essere sostituita con /

**header=TRUE /FALSE.**

Se header=TRUE specifica che la prima riga del file corrisponde al nome delle variabili.

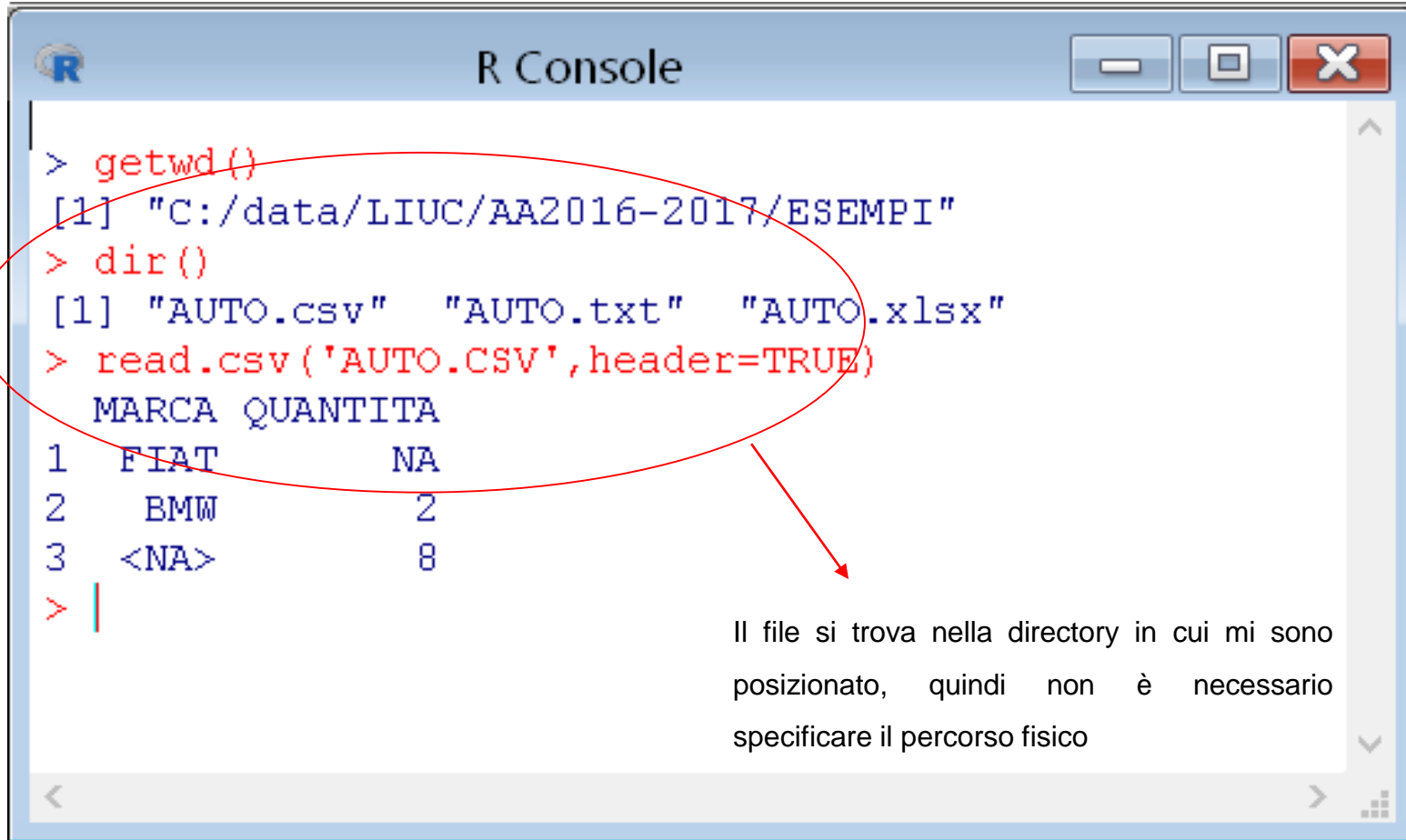
**Sep=','.**

Nel caso in cui si utilizzi la funzione read.table, bisogna specificare il tipo di separatore tra le variabili, che nel caso del CSV è la virgola



# Importazione file esterno (2/3)

## ESEMPIO DI LETTURA DATI IN R



```
> getwd()
[1] "C:/data/LIUC/AA2016-2017/ESEMPI"
> dir()
[1] "AUTO.csv" "AUTO.txt" "AUTO.xlsx"
> read.csv('AUTO.CSV',header=TRUE)
  MARCA QUANTITA
1  FIAT         NA
2  BMW          2
3 <NA>         8
> |
```

Il file si trova nella directory in cui mi sono posizionato, quindi non è necessario specificare il percorso fisico

# Importazione file esterno (3/3)

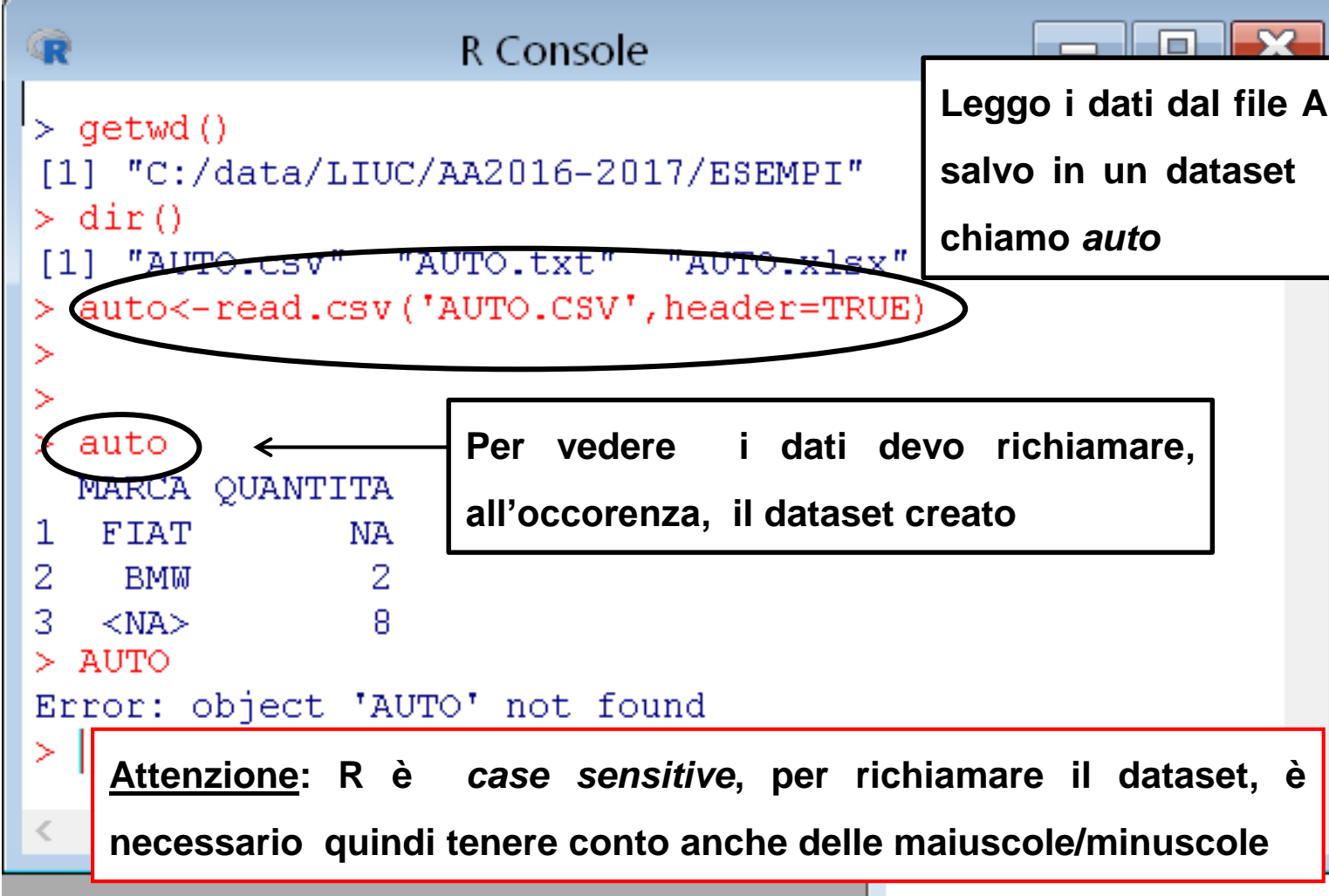
Nel modo appena descritto i dati però vengono semplicemente letti in R (viene visualizzato il file) ma non salvati, e quindi non sono utilizzabili per analisi successive.

Salvo il file in un oggetto chiamato **NOME\_DATASET** ← Leggo il file

```
NOME_DATASET ← read.csv('nome_file.csv', header=TRUE)
```

In questo modo creiamo un “oggetto”, con il nome da noi scelto, che potrà essere richiamato, durante la sessione R, ogni volta ce ne sia l'occorenza

# Importazione file esterno (3/3)



```
> getwd()
[1] "C:/data/LIUC/AA2016-2017/ESEMPI"
> dir()
[1] "AUTO.csv" "AUTO.txt" "AUTO.xlsx"
> auto<-read.csv('AUTO.CSV',header=TRUE)
>
>
> auto
  MARCA QUANTITA
1  FIAT         NA
2  BMW          2
3 <NA>         8
> AUTO
Error: object 'AUTO' not found
>
```

Leggo i dati dal file AUTO.CSV e li salvo in un dataset (oggetto) che chiamo *auto*

Per vedere i dati devo richiamare, all'occorrenza, il dataset creato

**Attenzione:** R è *case sensitive*, per richiamare il dataset, è necessario quindi tenere conto anche delle maiuscole/minuscole

# Come utilizzare una variabile(1/2)

Abbiamo visto come importare i dati nell'ambiente R.  
Come utilizzare ora le singole variabili?

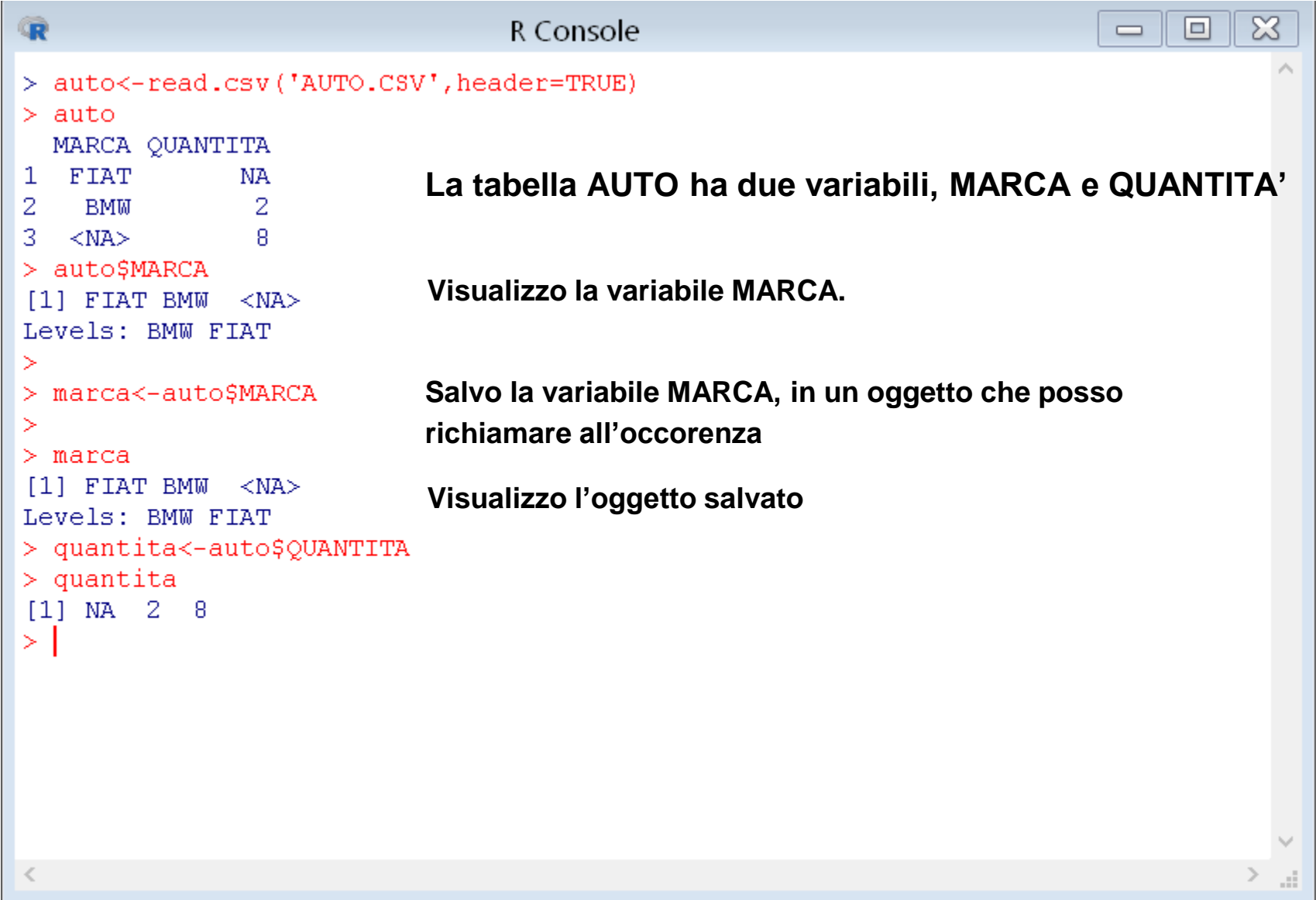
Per selezionare una variabile da una tabella è necessario scrivere il nome dataset seguito dal simbolo \$ e dal nome della variabile da selezionare

***Nome\_dataset\$nome\_variabile***

Anche in questo caso è possibile salvare la singola variabile in un oggetto, che è possibile richiamare all'occorrenza

***NEW\_NOME <- Nome\_dataset\$nome\_variabile***

# Come utilizzare una variabile(2/2)



The screenshot shows an R Console window with the following content:

```
> auto<-read.csv('AUTO.CSV',header=TRUE)
> auto
  MARCA QUANTITA
1  FIAT         NA
2  BMW          2
3  <NA>         8
> auto$MARCA
[1] FIAT BMW <NA>
Levels: BMW FIAT
>
> marca<-auto$MARCA
>
> marca
[1] FIAT BMW <NA>
Levels: BMW FIAT
> quantita<-auto$QUANTITA
> quantita
[1] NA  2  8
> |
```

**La tabella AUTO ha due variabili, MARCA e QUANTITA'**

**Visualizzo la variabile MARCA.**

**Salvo la variabile MARCA, in un oggetto che posso richiamare all'occorrenza**

**Visualizzo l'oggetto salvato**



# STR – Sintassi generale

Procedura che consente di visualizzare informazioni relative alle variabili contenute nel dataset (nome, tipo, lunghezza...)

```
str(nomedataset)
```



# HEAD – Sintassi generale

Procedura che consente di visualizzare le prime 6 osservazioni (righe) del dataset, funzione molto utile quando si ha a che fare con matrici di dati molto grandi.

```
head(nomedataset)
```



# FIX – Sintassi generale

Procedura che consente di visualizzare in forma tabellare il dataset importato.

N.B.: si possono modificare i valori delle variabili, in questo caso si modificherà l'oggetto importato in R e non il file originale.

```
fix(nomedataset)
```



# R: Output

In R i risultati di un'analisi vengono visualizzati nel prompt, ovvero dove si eseguono le funzioni.

```
> telefonia<-read.csv('telefonia.csv',header=TRUE)
> summary(telefonia)
```

n_questionario	tecnologia	marca	motivo_utilizzo_1
Min. : 1.00	Gsm :93	Nokia :103	Famigliari: 6
1st Qu.: 60.75	Non so :55	Motorola : 52	Lavoro : 28
Median :122.50	Palmare: 2	Samsung : 40	Partner : 1
Mean :121.78	Umts :86	Lg : 13	Piacere : 58
3rd Qu.:181.50		Sony Ericsson: 12	Studio :143
Max. :243.00		Altro : 6	
		(Other) : 10	

Comando di R

Output di R



Visualizzazione di default



# Esempio - telefonia

Importare il dataset telefonia.

```
> telefonia<-read.csv('telefonia.csv',header=TRUE)
```

Visualizzo le informazioni del dataset telefonia

```
> str(telefonia)
```

Visualizzo le prime 6 righe del dataset

```
> head(telefonia)
```

Visualizzo in forma tabellare l'intero dataset

```
> fix(telefonia)
```

# Esempio - telefonia - str

Visualizzo le informazioni del dataset telefonia

> str(telefonia)

```
> telefonia<-read.csv('telefonia.csv',header=TRUE)
> str(telefonia)
'data.frame': 236 obs. of 133 variables:
 $ n_questionario      : int  238 6 14 44 68 74 75 87 107 190 ...
 $ tecnologia          : Factor w/ 4 levels "Gsm","Non so",...: 4 2 4 2 4 2 1 4 4 2 ...
 $ marca               : Factor w/ 9 levels "Altro","Lg","Motorola",...: 5 5 5 5 5 3 3 7 3 5 $
 $ motivo_utilizzo_1   : Factor w/ 5 levels "Famigliari","Lavoro",...: 5 5 5 1 4 5 5 4 5 4 ...
 $ motivo_utilizzo_2   : Factor w/ 6 levels "", "Altro", "Famigliari",...: 5 5 5 1 4 3 5 4 5 1 $
 $ motivo_utilizzo_3   : Factor w/ 5 levels "", "Altro", "Famigliari",...: 3 3 3 1 1 4 4 1 3 1 $
 $ motivo_utilizzo_cell_1 : Factor w/ 5 levels "Famigliari","Lavoro",...: 5 5 5 5 4 5 4 4 4 4 ...
 $ motivo_utilizzo_cell_2 : Factor w/ 6 levels "", "Altro", "Famigliari",...: 5 5 5 5 4 3 4 4 3 1 $
 $ motivo_utilizzo_cell_3 : Factor w/ 5 levels "", "Altro", "Famigliari",...: 3 3 3 3 1 4 1 1 1 1 $
 $ operatore           : Factor w/ 4 levels "Tim","Tre","Vodafone",...: 3 3 3 3 3 4 3 3 2 3 .$
 $ operatore_cambio_2  : Factor w/ 4 levels "", "Influenzato da altri",...: 1 1 1 1 1 1 2 1 1 $
 $ operatore_scelta_1  : Factor w/ 10 levels "", "Altro", "Consigliato da conoscenti",...: 4 8 $
 $ operatore_scelta_2  : Factor w/ 9 levels "", "Altro", "Consigliato da conoscenti",...: 8 1 8$
 $ operatore_scelta_3  : Factor w/ 8 levels "", "Altro", "Consigliato da conoscenti",...: 6 1 4$
 $ internet            : Factor w/ 6 levels "1 volta/giorno",...: 4 4 4 4 4 4 4 4 5 4 ...
 $ piano_2             : Factor w/ 5 levels "", "Consiglio conoscenti",...: 5 5 3 1 5 4 5 5 3 $
 $ piano_3             : Factor w/ 4 levels ">2 volte","1 volta",...: 4 4 3 4 4 4 2 4 4 2 ...
```



# Esempio - telefonia - head

Visualizzo le prime 6 osservazioni delle variabili

> head(telefonia)

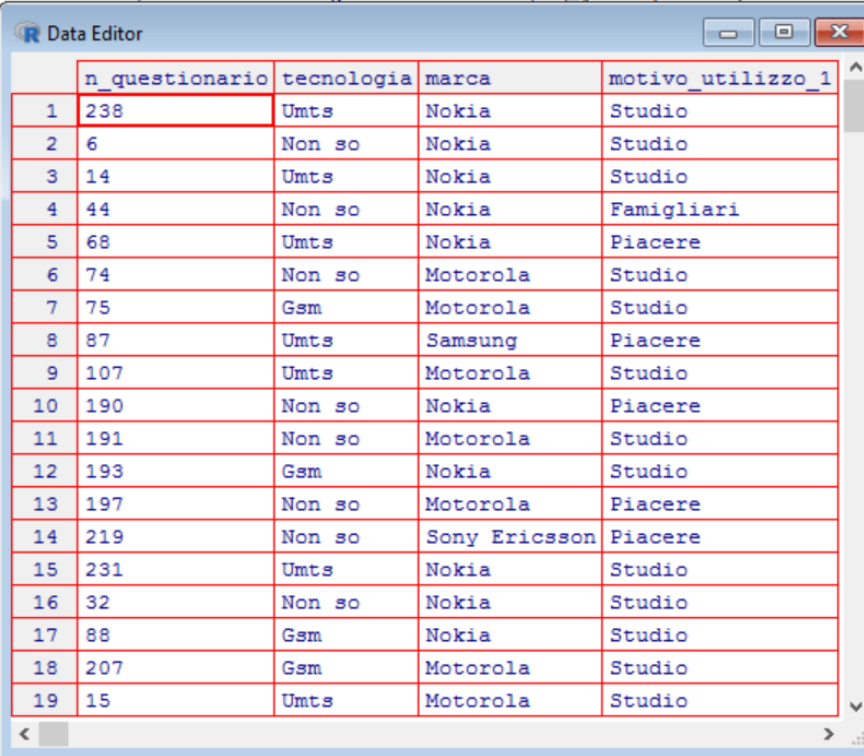
```
> head(telefonia)
  n_questionario tecnologia      marca motivo_utilizzo_1  motivo_utilizzo_2 motivo_utilizzo_3
1             238      Umts      Nokia              Studio Piacere/Tempo libero      Familiari
2              6      Non so      Nokia              Studio Piacere/Tempo libero      Familiari
3             14      Umts      Nokia              Studio Piacere/Tempo libero      Familiari
4             44      Non so      Nokia              Familiari
5             68      Umts      Nokia              Piacere              Partner
6             74      Non so Motorola              Studio              Familiari              Partner
motivo_utilizzo_cell_1 motivo_utilizzo_cell_2 motivo_utilizzo_cell_3 operatore
1              Studio Piacere/Tempo libero      Familiari Vodafone
2              Studio Piacere/Tempo libero      Familiari Vodafone
3              Studio Piacere/Tempo libero      Familiari Vodafone
4              Studio Piacere/Tempo libero      Familiari Vodafone
5              Piacere              Partner              Vodafone
6              Studio              Familiari              Partner      Wind
```



# Esempio - telefonia - fix

Visualizzo in forma tabellare il dataset

> `fix(telefonia)`



The screenshot shows the R Data Editor window displaying a dataset with 19 rows and 5 columns. The columns are labeled: n\_questionario, tecnologia, marca, and motivo\_utilizzo\_1. The data is as follows:

	n_questionario	tecnologia	marca	motivo_utilizzo_1
1	238	Umts	Nokia	Studio
2	6	Non so	Nokia	Studio
3	14	Umts	Nokia	Studio
4	44	Non so	Nokia	Famigliari
5	68	Umts	Nokia	Piacere
6	74	Non so	Motorola	Studio
7	75	Gsm	Motorola	Studio
8	87	Umts	Samsung	Piacere
9	107	Umts	Motorola	Studio
10	190	Non so	Nokia	Piacere
11	191	Non so	Motorola	Studio
12	193	Gsm	Nokia	Studio
13	197	Non so	Motorola	Piacere
14	219	Non so	Sony Ericsson	Piacere
15	231	Umts	Nokia	Studio
16	32	Non so	Nokia	Studio
17	88	Gsm	Nokia	Studio
18	207	Gsm	Motorola	Studio
19	15	Umts	Motorola	Studio





# Ordinare un dataset– Sintassi generale

Come ordinare una **singola variabile**?

```
sort(nome_variabile, decreasing=FALSE)
```

Se viene specificata l'opzione *decreasing=FALSE*, le osservazioni sono ordinate per valori crescenti della variabile, altrimenti se *decreasing=TRUE* per valori decrescenti.

Come ordinare **un dataset** rispetto una variabile?

```
Nome_dataset[order(nome_dataset$variabile,  
decreasing=FALSE), ]
```

Se viene specificata l'opzione *decreasing=FALSE*, le osservazioni sono ordinate per valori crescenti della variabile, altrimenti se *decreasing=TRUE* per valori decrescenti.



# Ordinare un dataset : Esempio

```
>
> auto$MARCA
[1] FIAT BMW <NA>
Levels: BMW FIAT
> sort(auto$MARCA,descending=FALSE)
[1] BMW FIAT
Levels: BMW FIAT
>
> auto
  MARCA QUANTITA
1  FIAT         NA
2   BMW          2
3 <NA>          8
> datiordinati<-auto[order(auto$MARCA,decreasing=FALSE),]
> datiordinati
  MARCA QUANTITA
2   BMW          2
1  FIAT         NA
3 <NA>          8
> |
```

**Ordino la variabile  
MARCA**


**Ordino il dataset  
rispetto alla variabile  
MARCA**



# Ordinare un dataset : Esempio - telefonia


## Ordino la variabile OPERATORE

```
> operatore<-telefonia$operatore
> sort(operatore,decreasing=TRUE)
 [1] Wind      Wind      Wind      Wind      Wind      Wind      Wind      Wind      Wind      Wind
[11] Wind      Wind      Wind      Wind      Wind      Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone
[21] Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone
[31] Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone
[41] Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone
[51] Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone Vodafone
```



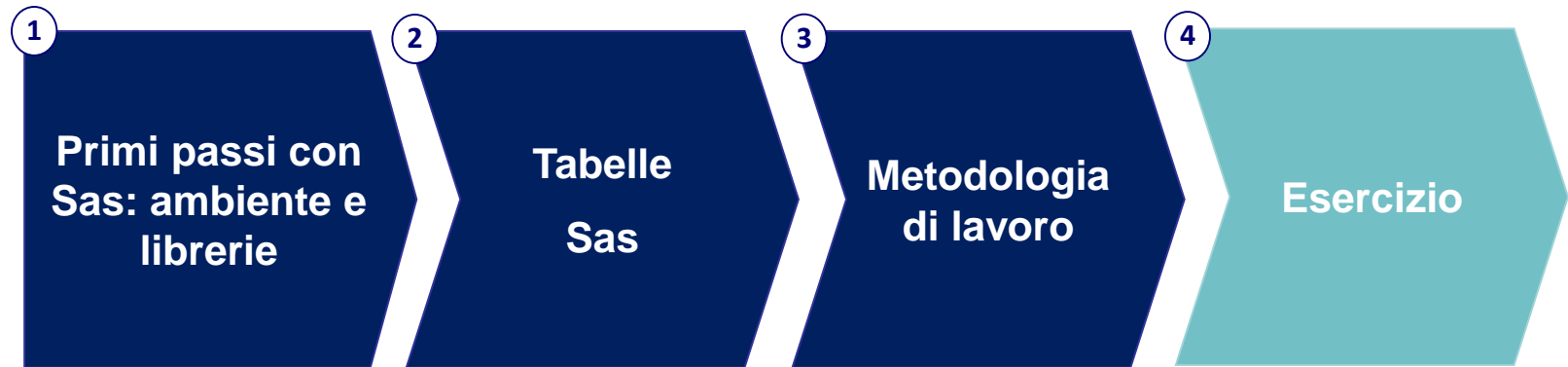
## Ordino il dataset rispetto alla variabile PROFESSIONE

```
> telefonia[order(telefonia$professione,decreasing=FALSE), ]
```



# Metodi Quantitativi per Economia, Finanza e Management

**Obiettivi di questa esercitazione:**



# Esercizi

Il file Excel DENTI.CSV contiene dati sul consumo di dentifricio (di marca A e di marca B).

1. **Allocare la DIRECTORY DI LAVORO** (che punta alla cartella che contiene il file DENTI.CSV).
2. **Importare in R** la tabella DENTI.CSV e salvarla in un oggetto col nome DENTI\_NEW.
3. Visualizzare le **caratteristiche delle variabili** tramite opportuna procedura:
  - Quante osservazioni contiene il dataset?
  - Quante variabili contiene il dataset?
  - Di che tipo è la variabile **REGIONE, secondo R?**
4. Ordinare le osservazioni della tabella in senso decrescente per consumo totale di dentifricio (CONSTOT).

