

L' ambiente grafico di R

- Il linguaggio R è dotato di un ambiente grafico potente e versatile
- E' semplice produrre grafici per l' analisi esplorativa dei dati
- Si possono facilmente generare grafici di elevata qualità utilizzabili per pubblicazioni
- L' ambiente grafico di R può generare grafici in diversi formati:
 - Display in finestre grafiche (Linux, Windows e Macintosh)
 - postscript
 - pdf (Adobe Portable Document Format)
 - jpeg (JPEG bitmap)
 - png (PNG bitmap, simile a GIF)
 - wmf (Windows Metafile)

Plot

- E' una funzione generica di R: il tipo di grafico generato dipende dal tipo o classe del suo argomento:
- **plot**(x,y) : se x e y sono vettori produce un grafico delle coppie (x_i, y_i)
- **plot**(X): se X è una matrice a due colonne produce il plot di una colonna rispetto all'altra
- **plot**(x): se x è un vettore produce un grafico dei valori del vettore rispetto agli indici
- **plot** (df): se df è un dataframe, produce i grafici delle distribuzioni delle variabili contenute nel data frame
- **plot** dispone di diversi argomenti (si veda l'help)

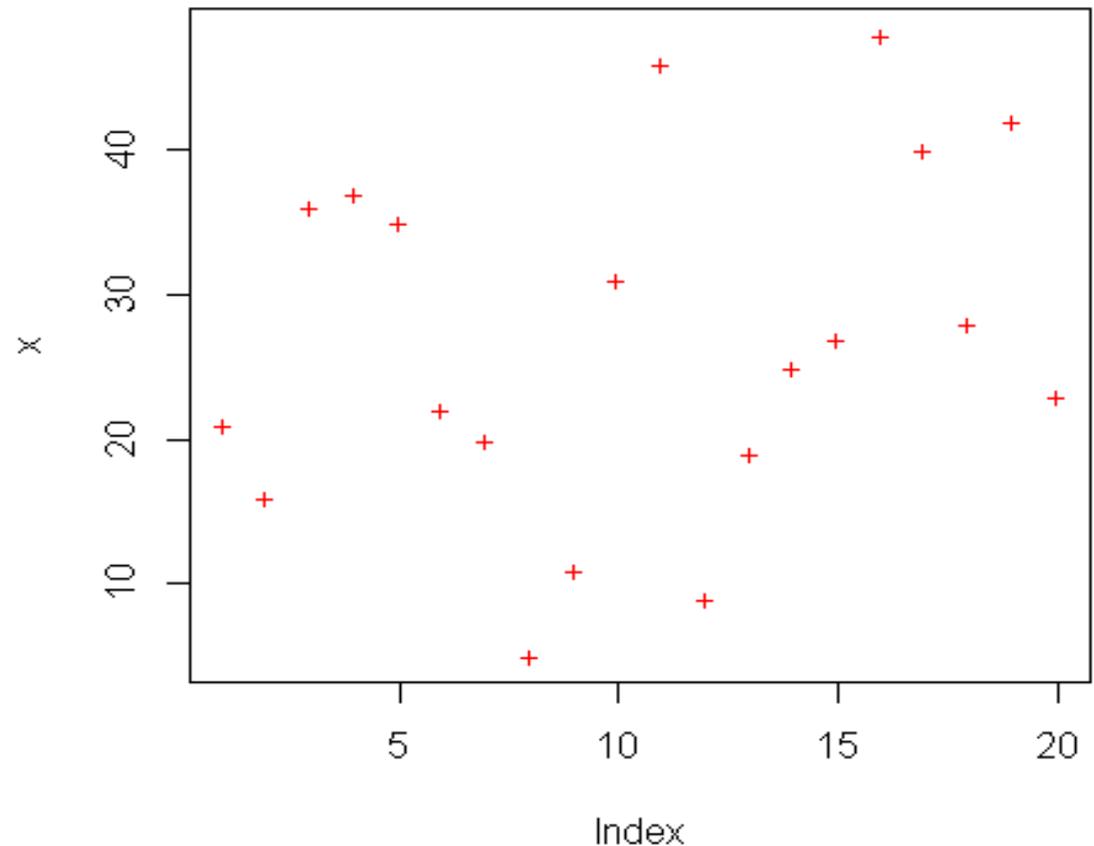
Plot di un vettore numerico: grafico di dispersione

```
> x <- sample(1:50, 20)  
> plot(x,col="2",pch="+")
```

Esistono diversi parametri
grafici:

Es.

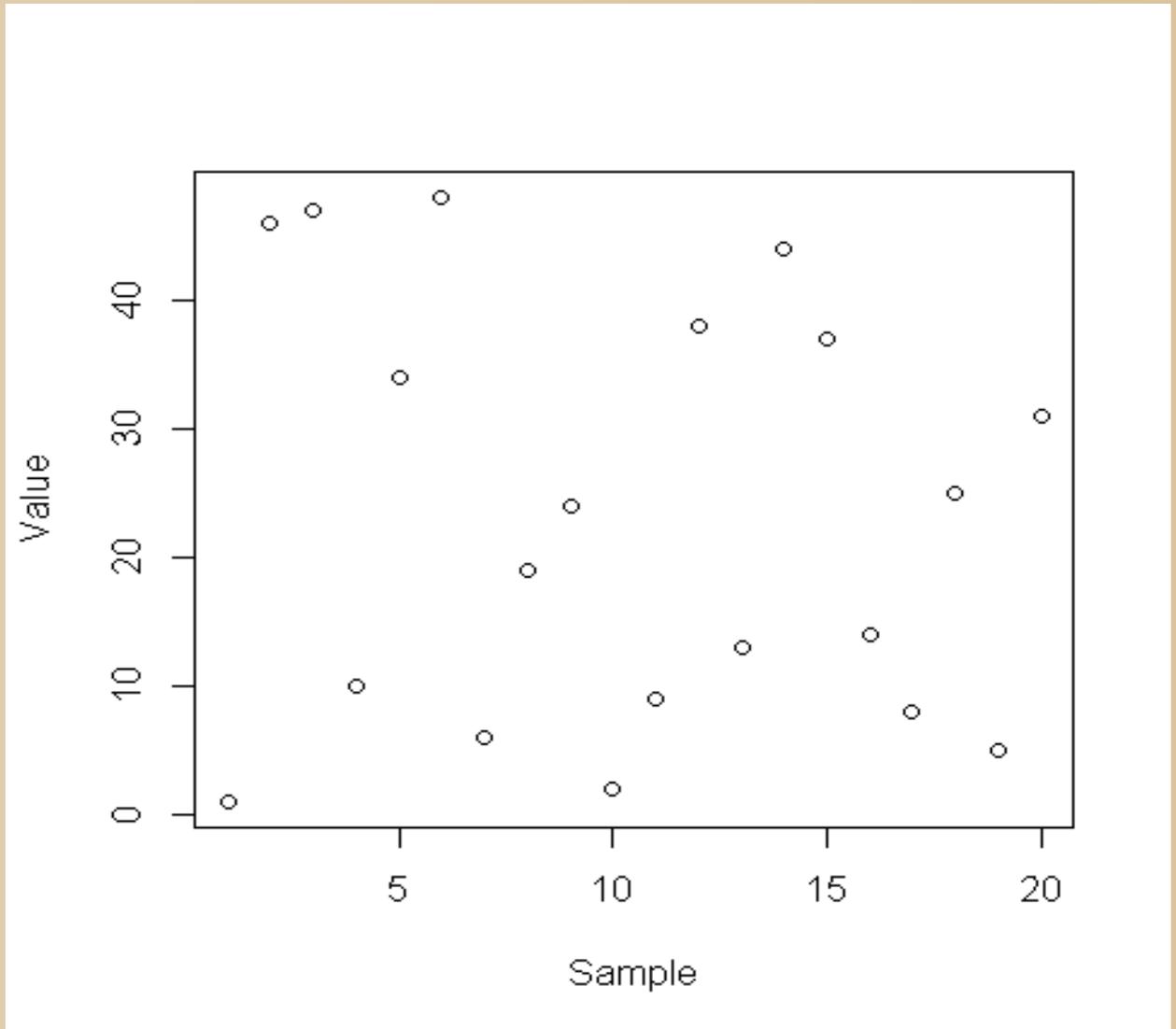
- pch : simbolo usato
 - col : colore dei punti/linee
- Etc.



plot : argomenti xlab - ylab

```
> x <- sample(1:50, 20)  
> plot(x, xlab="Sample",  
       ylab="Value")
```

xlab e **ylab** permettono di assegnare una etichetta agli assi

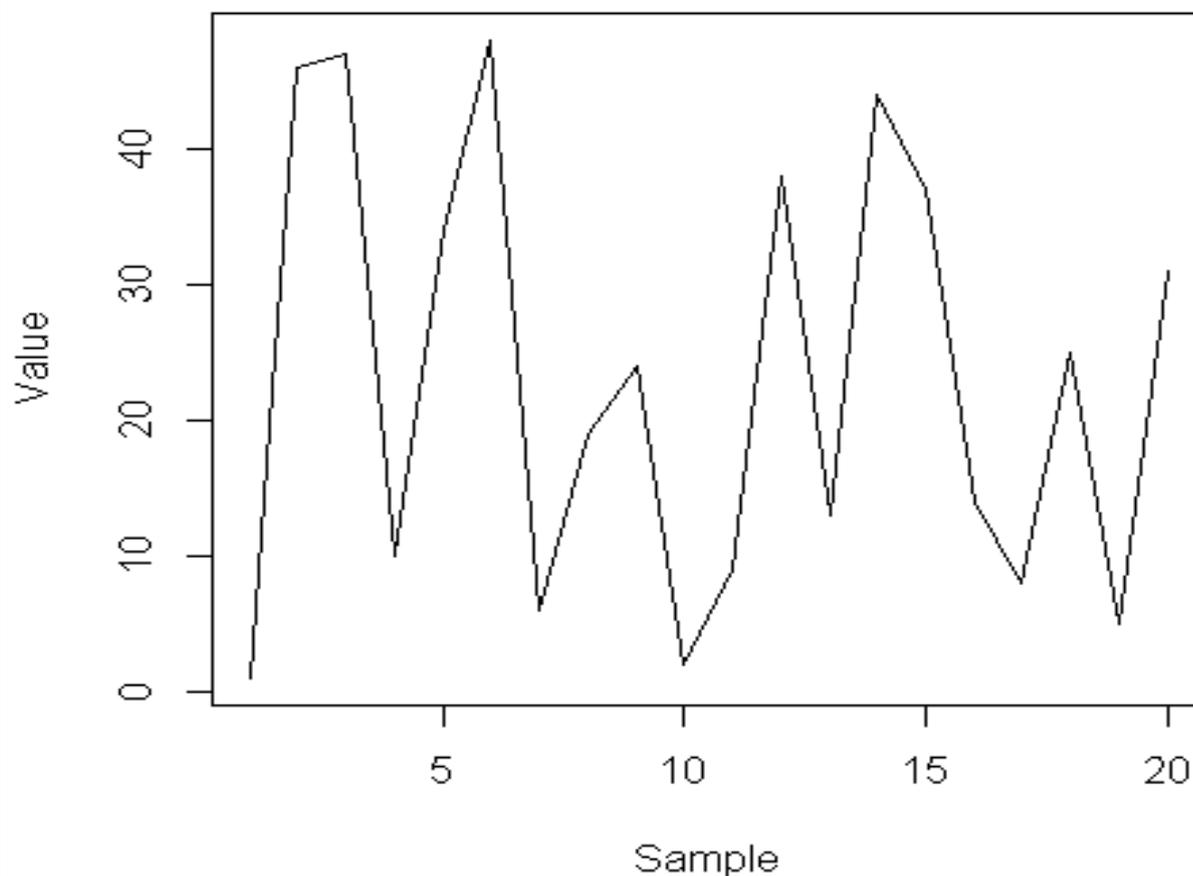


plot: argomento type

```
> x <- sample(1:50, 20)  
> plot(x, xlab="Sample",  
      ylab="Value", type="l")
```

type specifica il tipo di grafico: "l" sta per lines, "b" per both, linee e punti, etc.

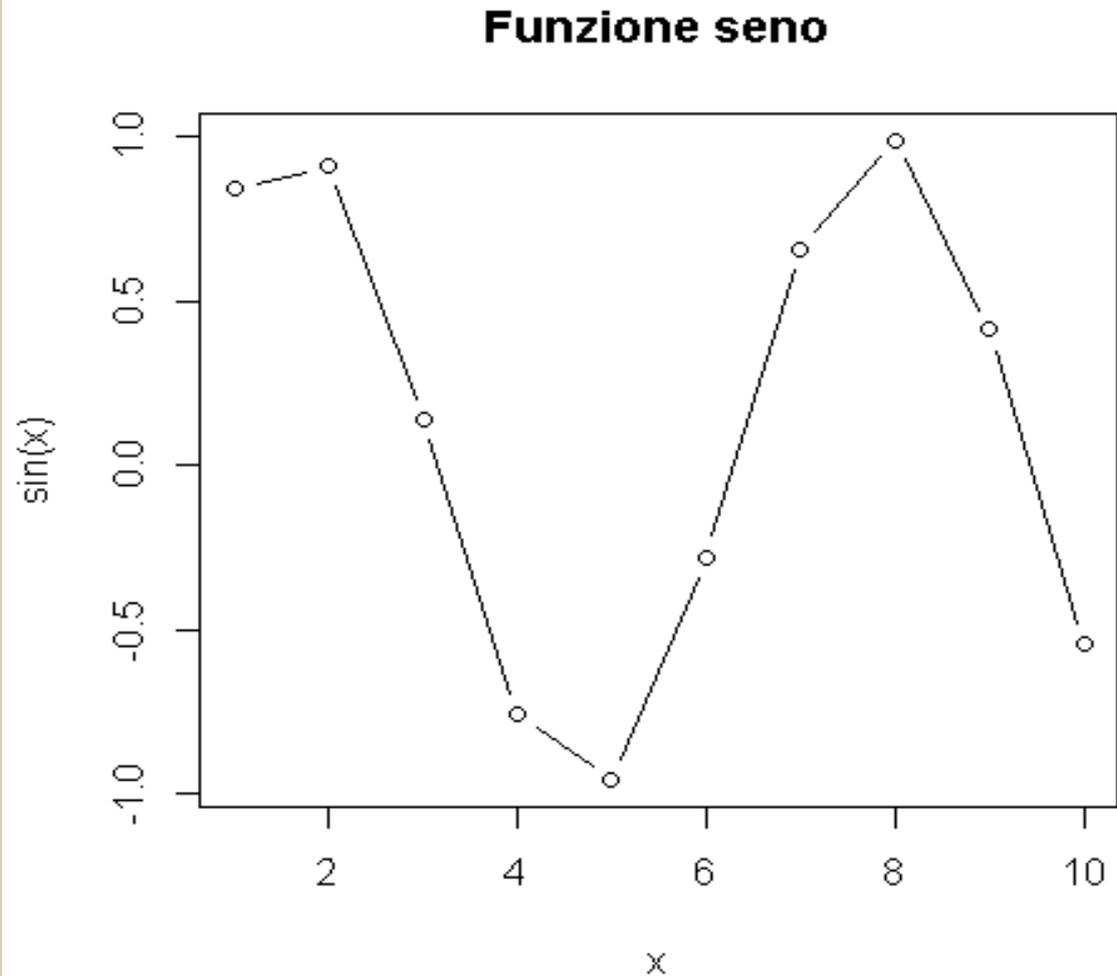
Vedi Help in linea, e provate



Plot di un vettore numerico rispetto ad un altro

```
x <- 1:10  
y <- sin(x)  
plot(x, y, xlab="x",  
      ylab="sin(x)",  
      type="b",  
      main="Funzione  
seno")
```

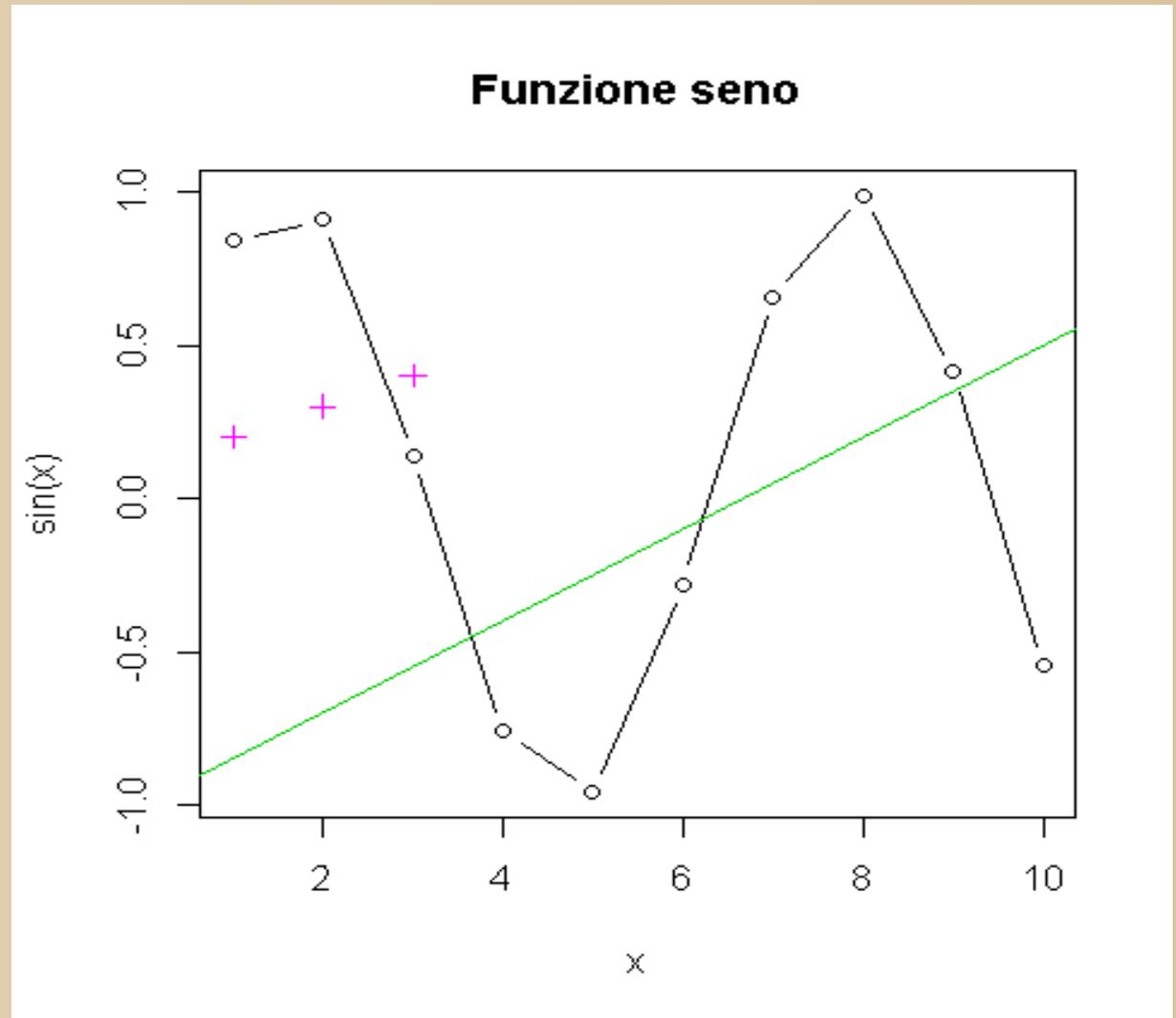
- L'argomento **main** permette di inserire un titolo
- Quello ottenuto è un grafico poligonale



Aggiungere una linea o dei punti

```
x <- 1:10
y <- sin(x)
plot(x, y, xlab="x",
      ylab="sin(x)",
      type="b",
      main="Funzione
      seno")
```

- **abline**($a=-1, b=0.15, col=3$)
 - a intercetta,
 - b coefficiente angolare
- **points**($c(1, 2, 3), c(0.2, 0.3, 0.4), pch=3, col=6$)



Salviamo una figura

- Apriamo un device grafico

```
> jpeg("primo_plot.jpg")
> x <- 1:10
> y <- sin(x)

> plot(x, y, xlab="x", ylab="sin(x)", type="l",
       main="Funzione seno")

> dev.off() # chiude il device grafico aperto
```

- Nella cartella attuale troveremo il file "primo_plot.jpg"
- Per gli altri formati grafici esistono i comandi corrispondenti (**postscript** per immagini .eps, **tiff** per i file .tiff, etc.)

Esercizi



ESERCIZIO 1.

- Inserire in una matrice i dati dei 30 lavoratori riportati nella tabella 2.12 a pag 48.
- riprodurre i grafici di dispersione delle figure 2.13 e 2.14 a pagina 49.

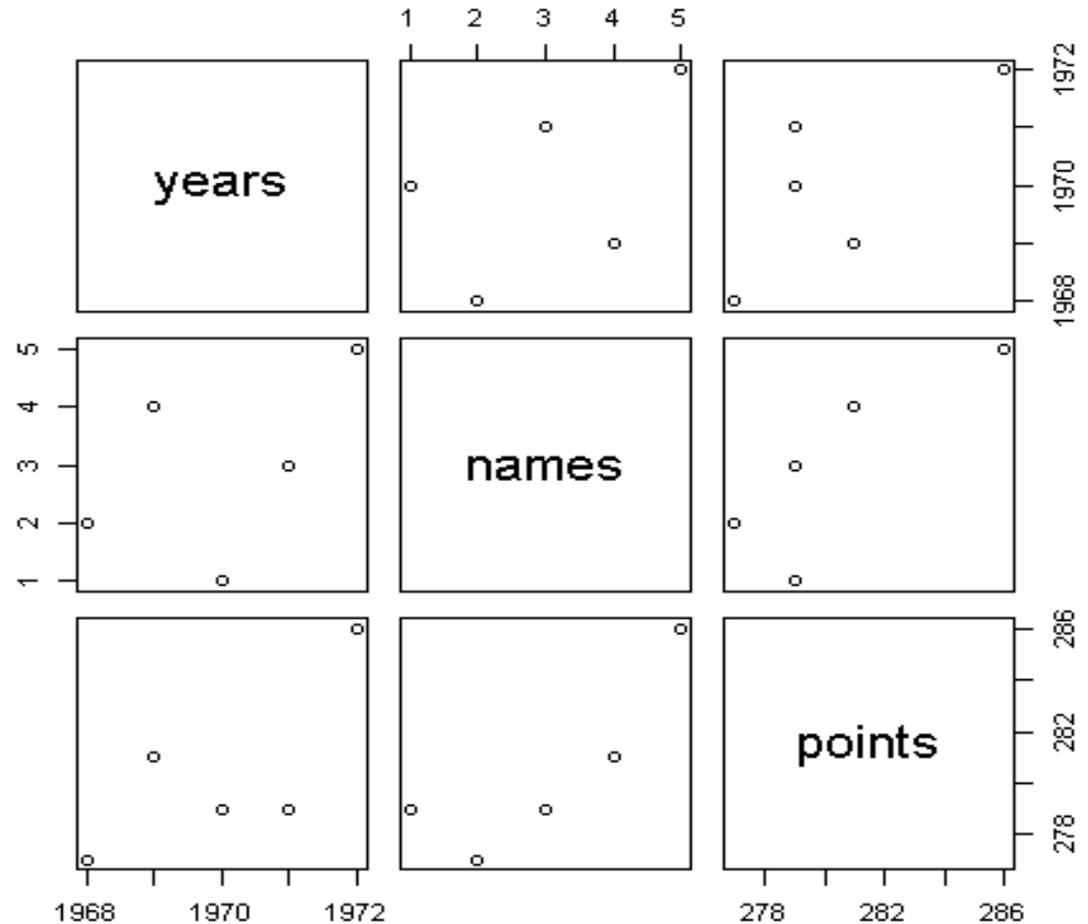
ESERCIZIO 2

- Rispondere ai punti (a) , (b) e (c) del problema 1 a pagina 50.



Plot di un dataframe

```
> years <- 1968:1972
> names <- c("Bob Goalby",
             "George Archer",
             "Billy Casper",
             "Charles Coody",
             "Jack Nicklaus")
> points <- c(277, 281, 279,
             279, 286);
> masters <- data.frame(years,
                        names, points)
> plot(masters)
```



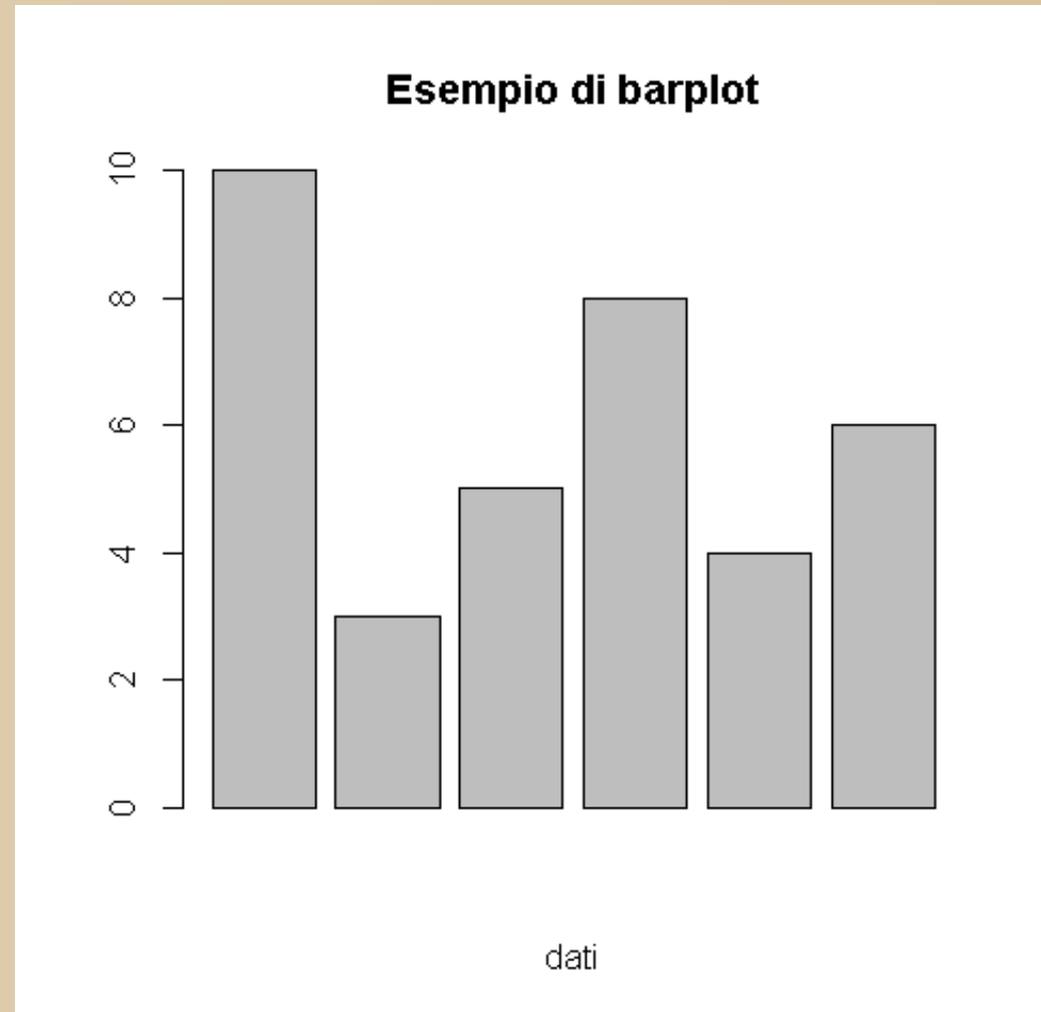
barplot

- **barplot(x)** genera un grafico a barre partendo dai valori del vettore x

```
> dati <- c(10, 3, 5, 8, 4, 6)
> barplot(dati, xlab="dati",
          main="Esempio di barplot")
```

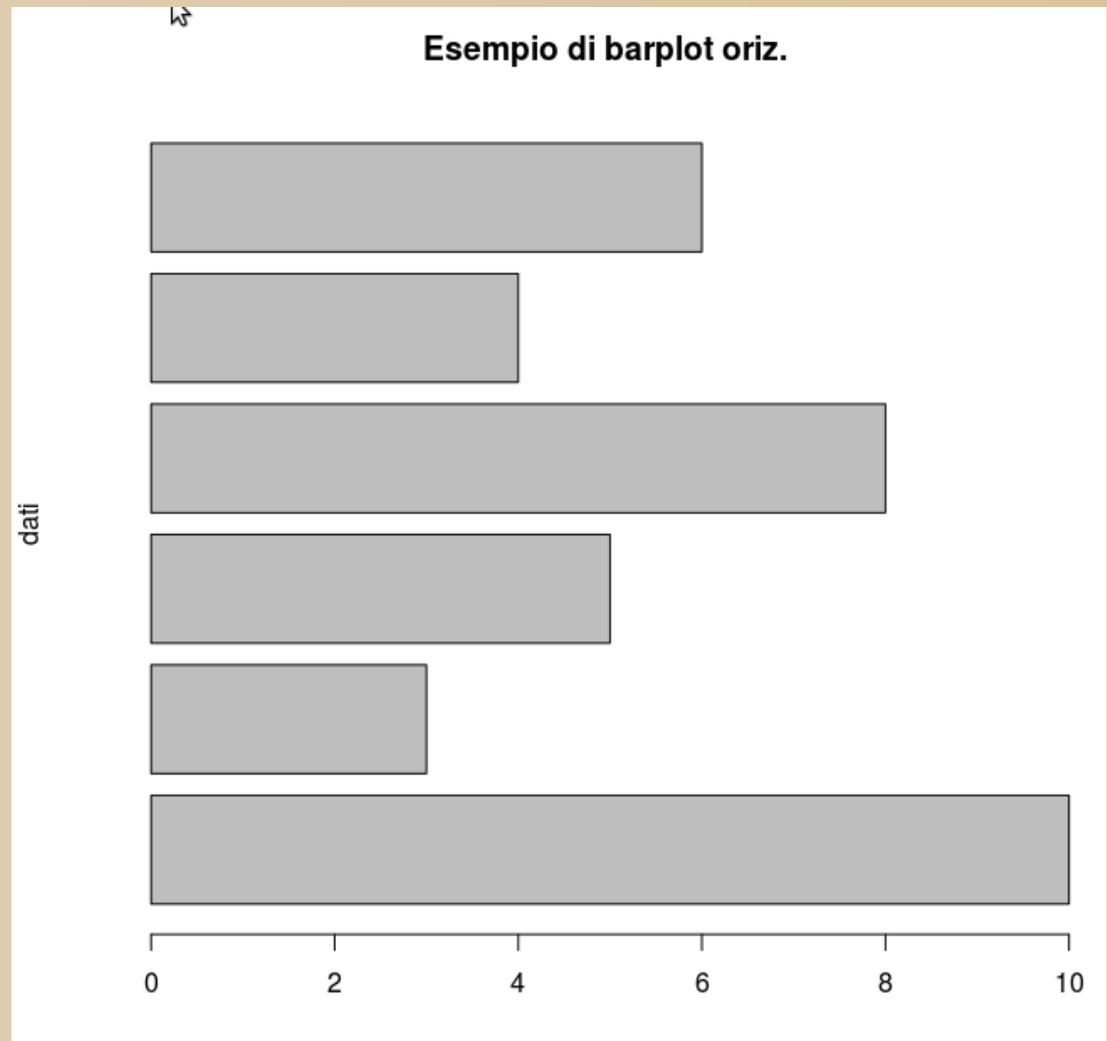
Es.

- width : per la larghezza delle barre
- horiz: per un grafico orizzontale
- Vedere l'Help in linea



barplot

```
> dati <- c(10, 3, 5, 8, 4, 6)  
> barplot(dati, ylab="dati",  
  main="Esempio di barplot",  
  horiz = TRUE)
```



Hist

- **hist** genera istogrammi utilizzando un vettore numerico.

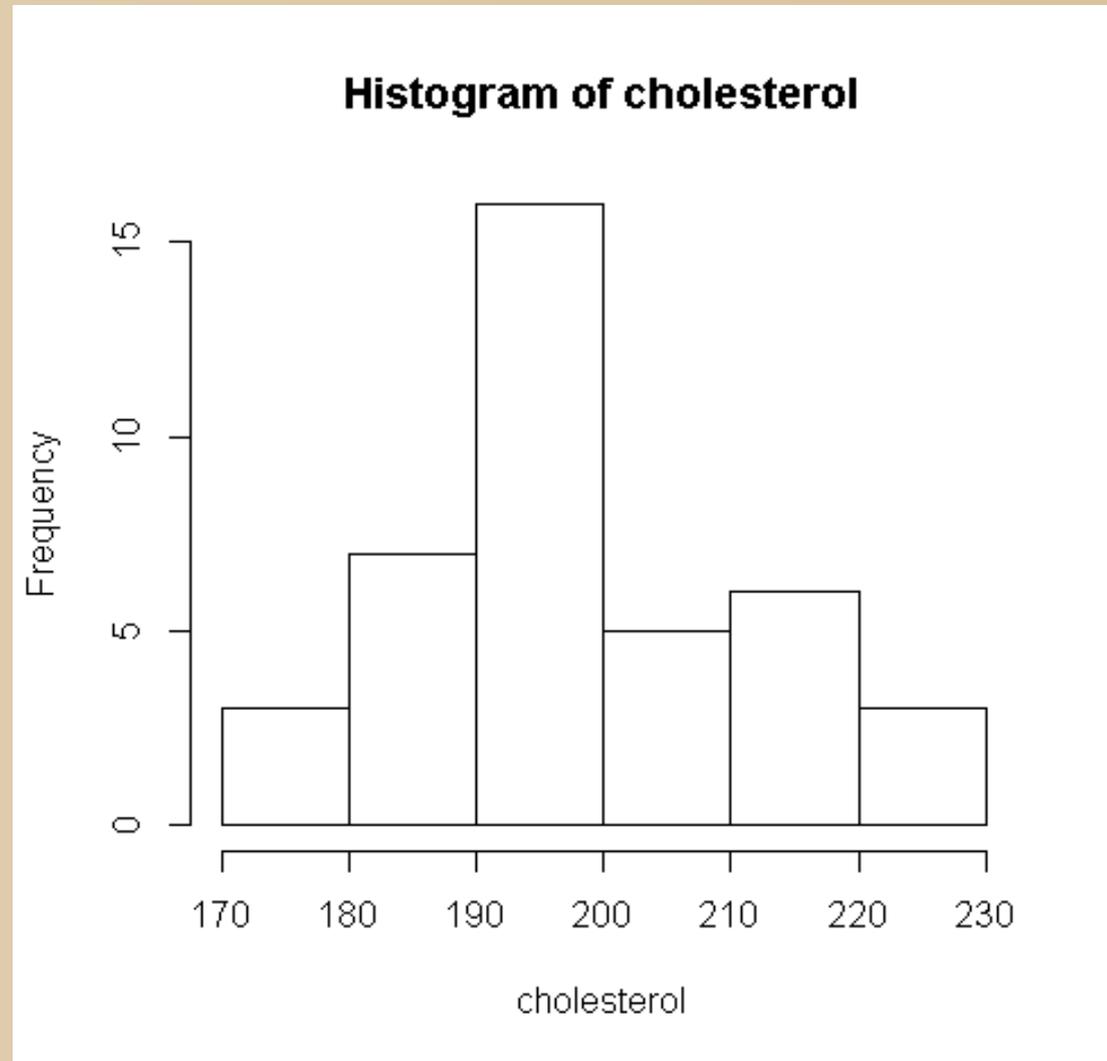
Esempi:

- `hist(x)`:
genera un istogramma utilizzando il vettore numerico `x`
- `hist(x, nclass=n)`:
genera un istogramma con un numero `n` di classi
- `hist(x, breaks=b, ...)`:
i punti di break degli intervalli dei valori di `x` che delimitano le classi sono esplicitamente elencati con il parametro **breaks**
- `hist(x, probability=TRUE)`
le colonne rappresentano frequenze relative invece che assolute

Hist

```
> cholesterol <- c(213, 174, 192,  
200, 187, 181, 216,  
206, 221, 212, 193,  
196, 220, 200, 199,  
178, 193, 205, 196,  
195, 191, 171, 221,  
204, 204, 183, 194,  
200, 183, 188, 193,  
211, 202, 213, 194,  
184, 191, 191, 183,  
227);
```

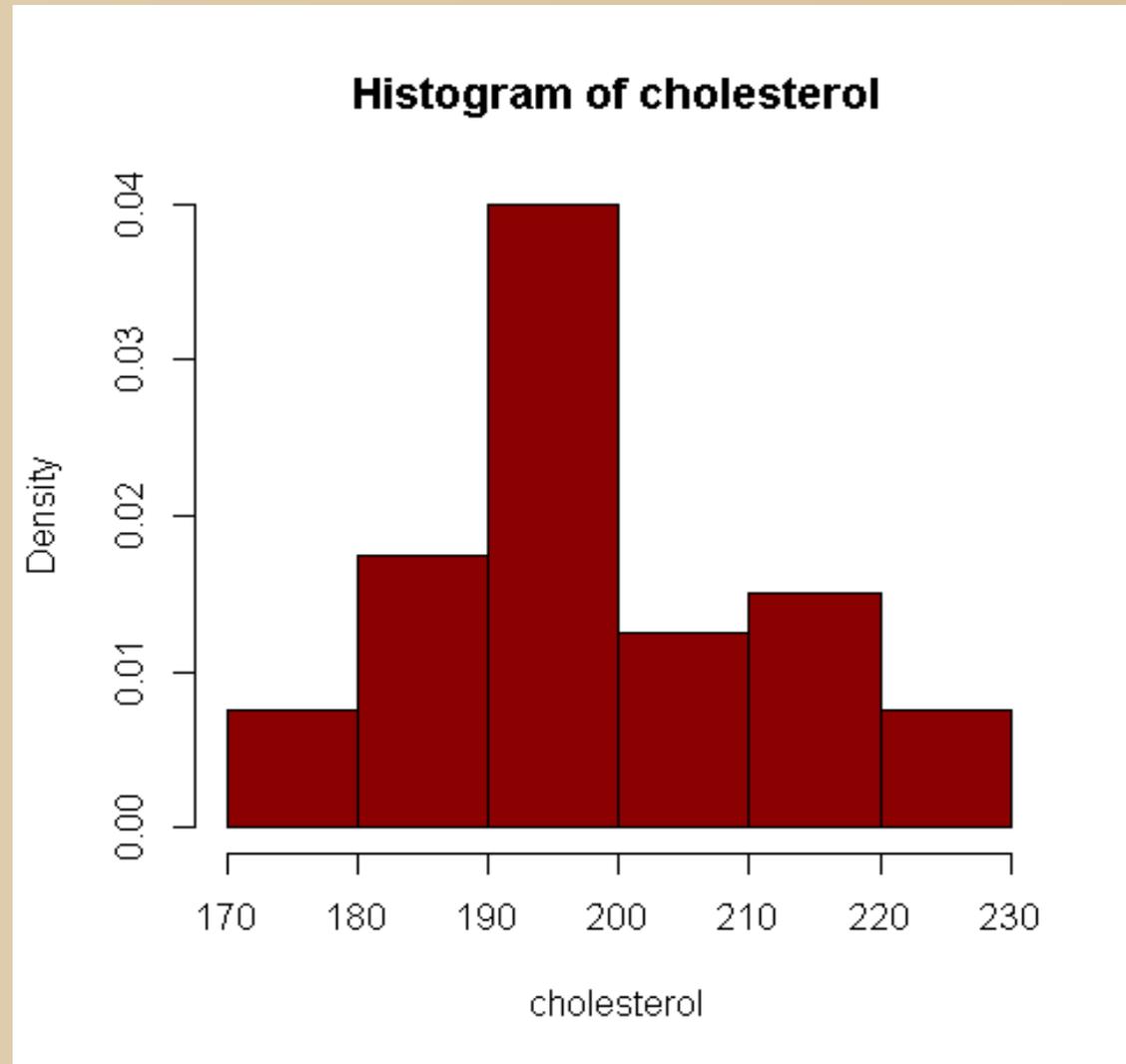
```
hist(cholesterol)
```



Hist

```
> cholesterol <- c(213, 174, 192,  
  200, 187, 181, 216,  
  206, 221, 212, 193,  
  196, 220, 200, 199,  
  178, 193, 205, 196,  
  195, 191, 171, 221,  
  204, 204, 183, 194,  
  200, 183, 188, 193,  
  211, 202, 213, 194,  
  184, 191, 191, 183,  
  227);
```

```
hist(cholesterol, probability =  
TRUE, col="darkred")
```



Hist: breaks

```
> cholesterol <- c(213, 174, 192,  
  200, 187, 181, 216,  
  206, 221, 212, 193,  
  196, 220, 200, 199,  
  178, 193, 205, 196,  
  195, 191, 171, 221,  
  204, 204, 183, 194,  
  200, 183, 188, 193,  
  211, 202, 213, 194,  
  184, 191, 191, 183,  
  227);
```

```
hist(cholesterol, breaks=  
  c(170, 185, 210, 230)))
```

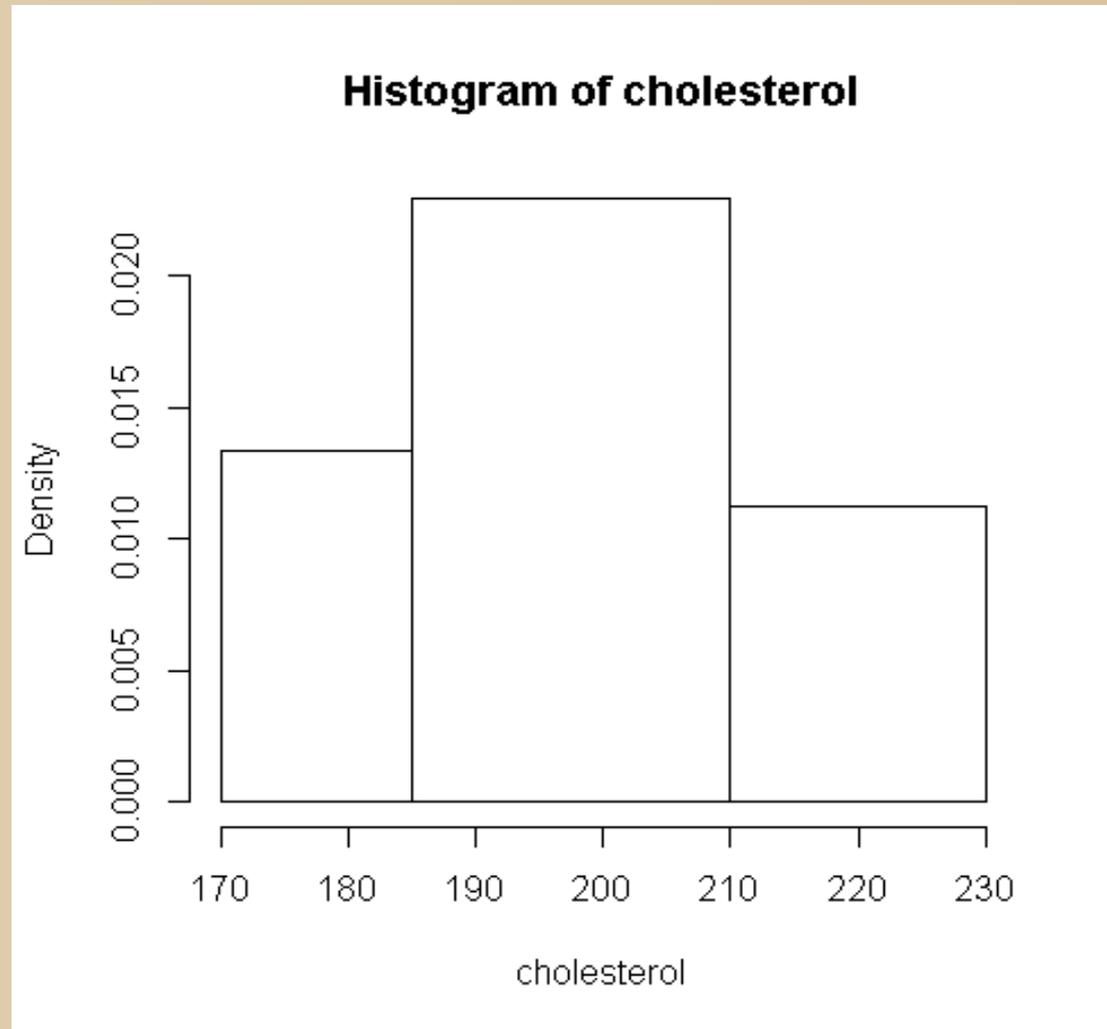


Diagramma a bastoncini

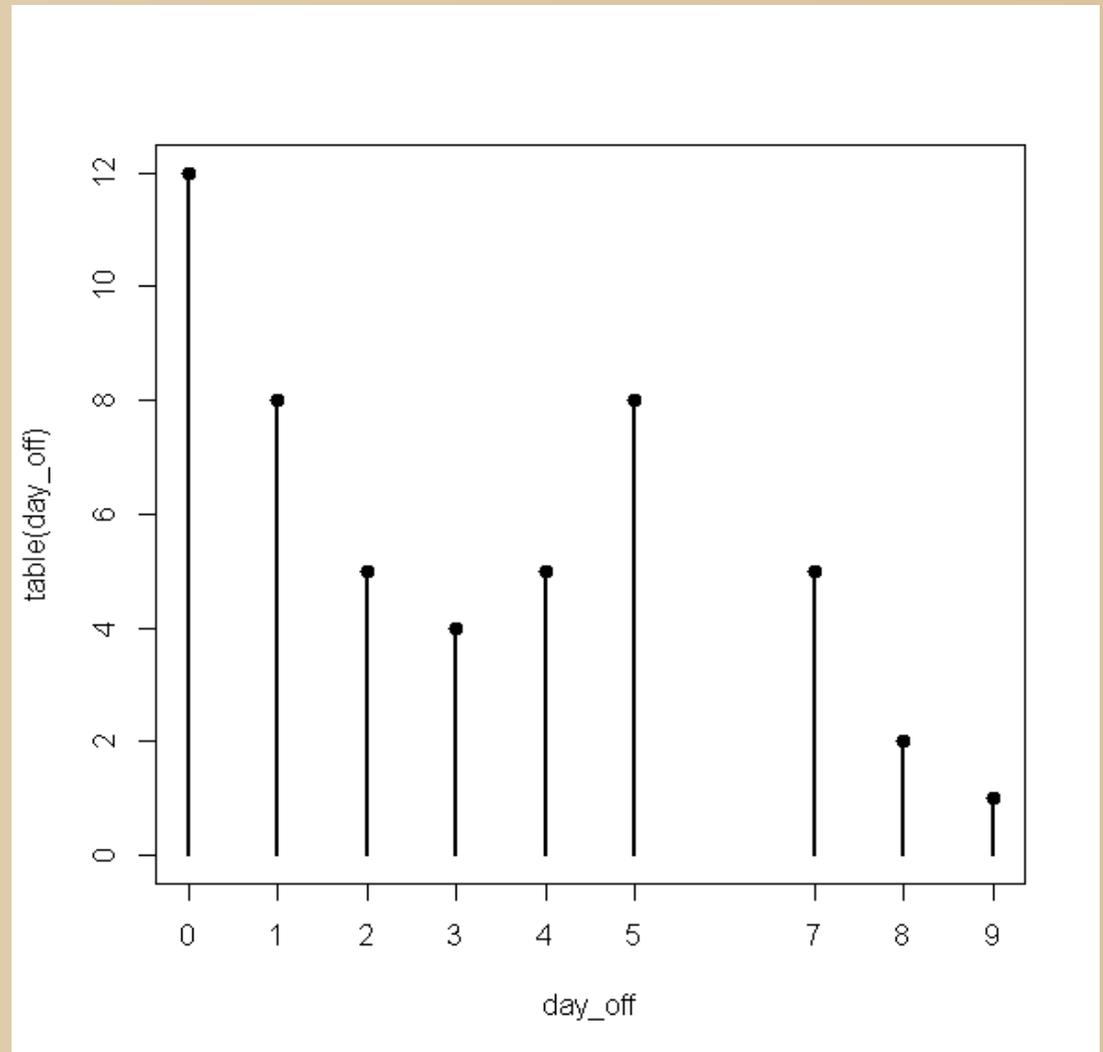
```
> day_off <- c(2, 2, 0, 0, 5, 8, 3,  
              4, 1, 0, 0, 7, 1, 7,  
              1, 5, 4, 0, 4, 0, 1,  
              8, 9, 7, 0, 1, 7, 2,  
              5, 5, 4, 3, 3, 0, 0,  
              2, 5, 1, 3, 0, 1, 0,  
              2, 4, 5, 0, 5, 7, 5, 1)
```

```
> freq <- table(day_off);
```

```
> plot(table(day_off), type="p",  
       bg="black", pch=21)
```

```
> lines(table(day_off)) # lines
```

necessita che sia stata già chiamata
la funzione plot



Diagrammi ramo foglia

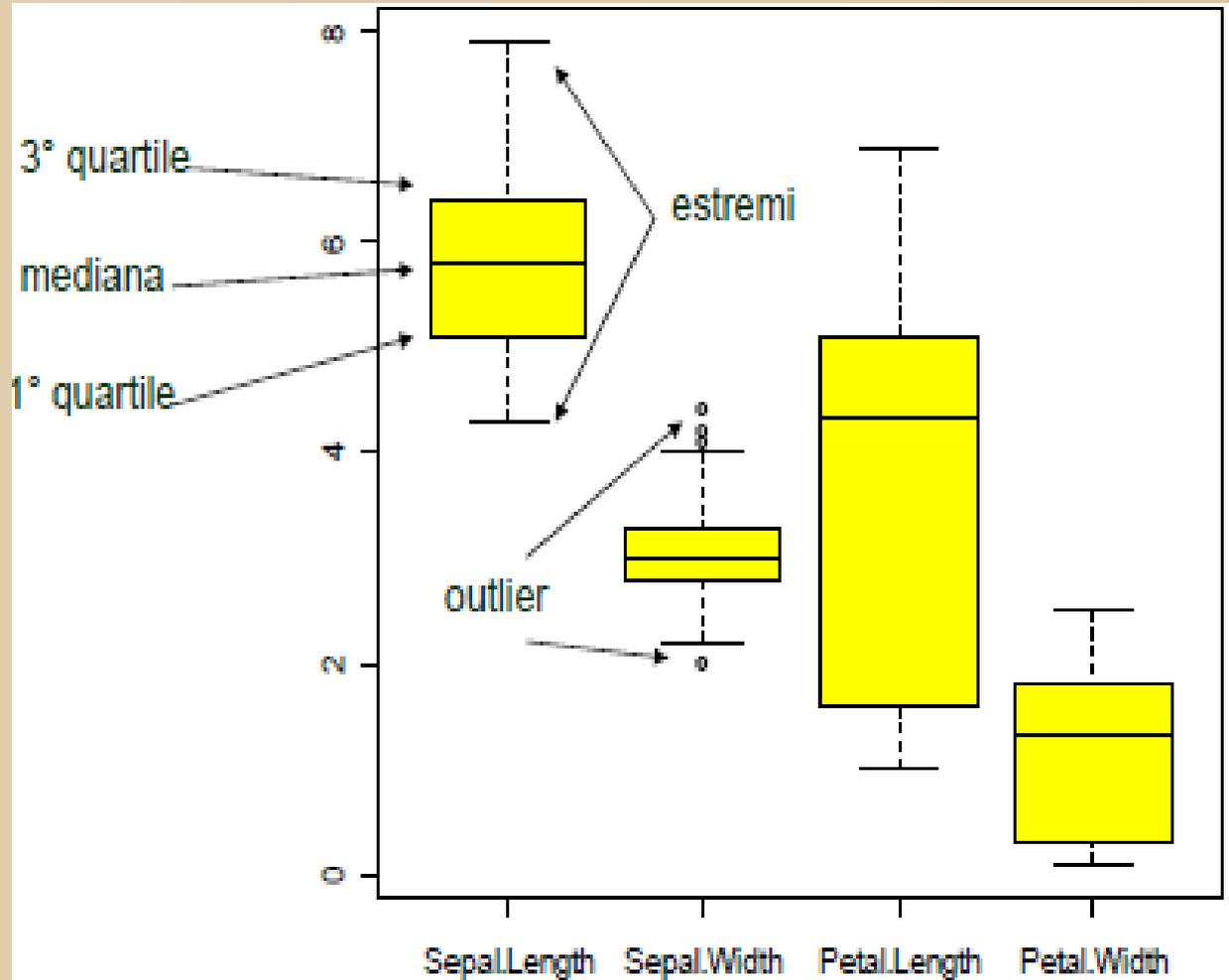
- In R i diagrammi ramo foglia si ottengono con il comando **stem**
 - Tramite l'argomento *scale* si può personalizzare la visualizzazione
- ```
> prova <- c(5, 27, 18, 11, 63, 39)
> stem(prova , scale=2)
```

The decimal point is 1 digit(s) to the right of the |

```
0 | 5
1 | 18
2 | 7
3 | 9
4 |
5 |
6 | 3
```

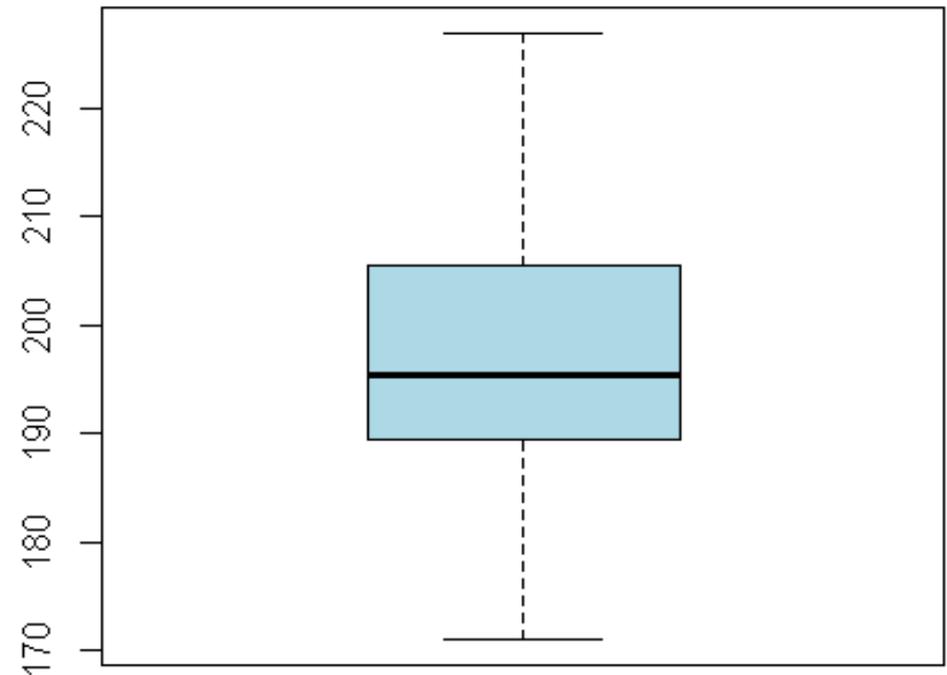
# Boxplot

- Forniscono una descrizione grafica sintetica di un insieme di dati utilizzando semplici statistiche
  - > `data(iris);`
  - > `boxplot(iris[,1:4], col="yellow")`



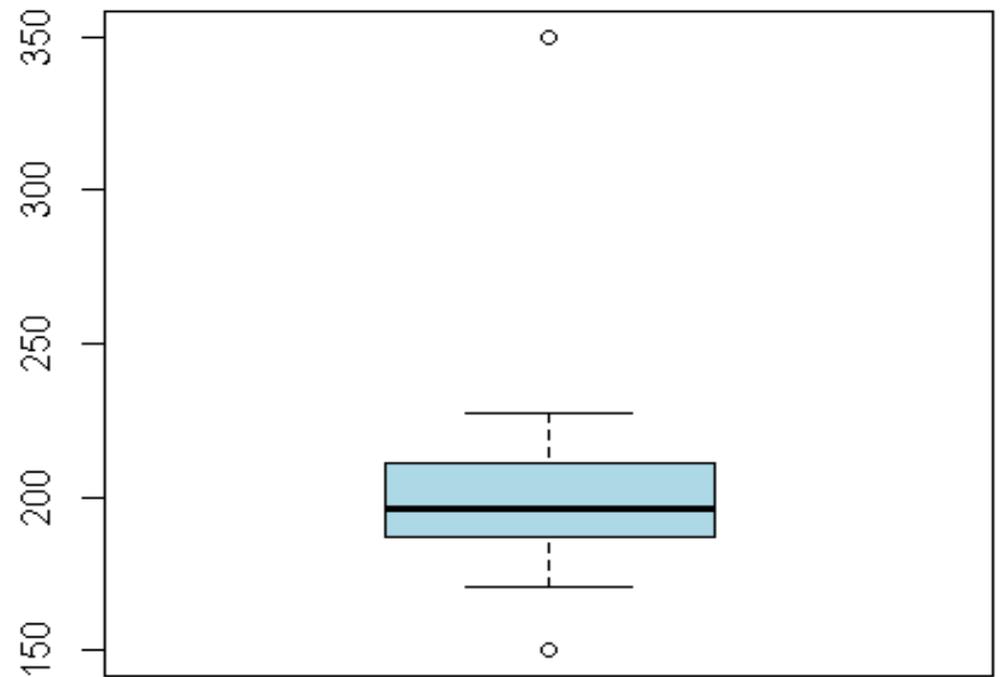
# Boxplot

```
> cholesterol <- c(213, 174, 192,
 200, 187, 181, 216,
 206, 221, 212, 193,
 196, 220, 200, 199,
 178, 193, 205, 196,
 195, 191, 171, 221,
 204, 204, 183, 194,
 200, 183, 188, 193,
 211, 202, 213, 194,
 184, 191, 191, 183,
 227);
> boxplot(cholesterol, col =
 "lightblue")
```



# Boxplot: outliers

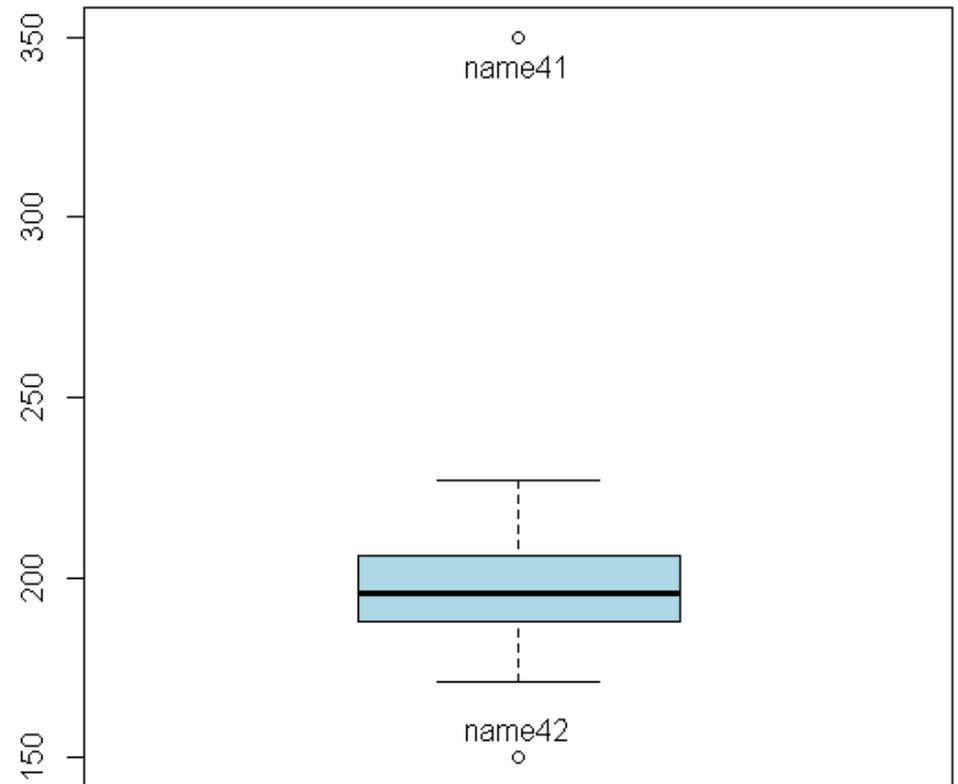
- Aggiungiamo degli outlier
- ```
> cholesterol <- c(cholesterol,  
                  350, 150);  
> boxplot(cholesterol, col =  
          "lightblue")
```



Boxplot: outliers

- **Identify** legge la posizione puntatore grafico quando il pulsante del mouse viene premuto, e cerca le coordinate del punto più vicino al puntatore, e ne stampa il nome

```
> n <- length(cholesterol)
> names(cholesterol) <-
paste0("name", 1:n)
> identify(rep(1, n), cholesterol,
           labels = names(cholesterol))
```
- Il primo vettore rappresenta le ascisse, tutti 1 perché i punti sono in verticale
- **labels** è l'etichetta associata ai punti

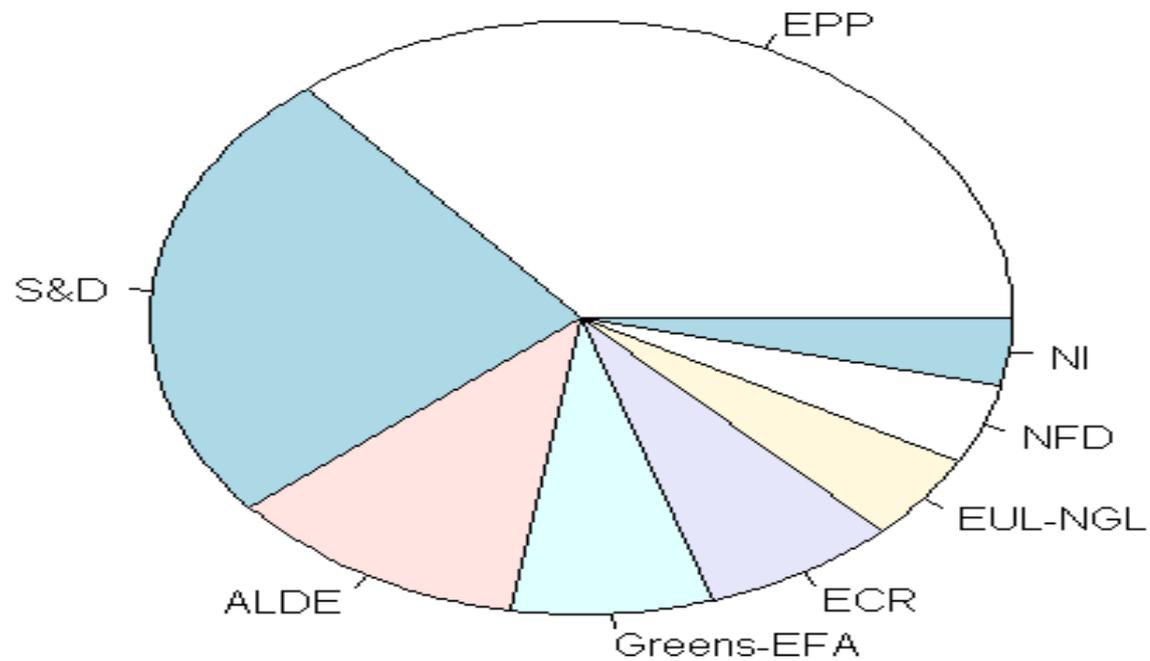


Grafici a torta

- In R i grafi a torta si creano con il comando `pie`
 - > `slices <- c(265, 184, 84, 55, 54, 35, 32, 27)`
 - > `lbls <- c("EPP", "S&D", "ALDE", "Greens-EFA", "ECR", "EUL-NGL", "NFD", "NI")`
 - > `pie(slices, labels = lbls, main="Grafico a torta dei partiti presenti all'EuroParlamento 2009")`

Grafici a torta

Grafico a torta dei partiti presenti all'EuroParlamento 2009

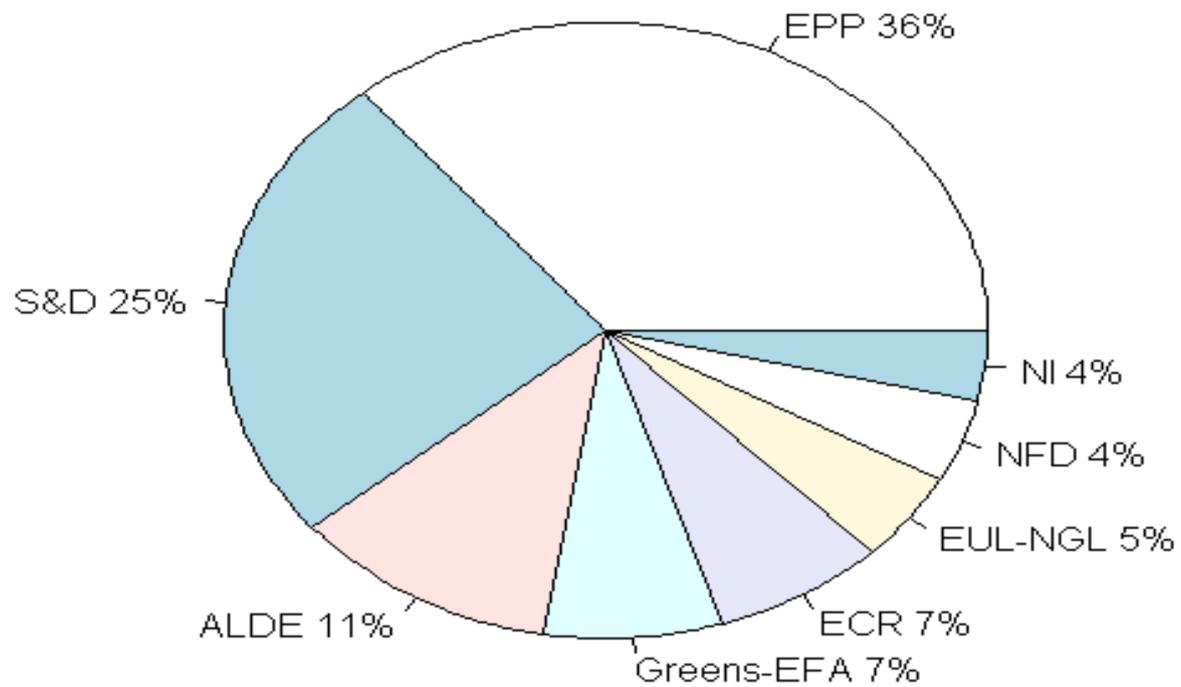


Grafici a torta

- **Può essere utile aggiungere le percentuali:**
- **slices <- c(265, 184, 84, 55, 54, 35, 32, 27)**
- **lbls <- c("EPP", "S&D", "ALDE", "Greens-EFA", "ECR", "EUL-NGL", "NFD", "NI")**
- **pct <- round(slices/sum(slices)*100) #calcolo delle percentuali**
- **lbls <- paste(lbls, pct) # aggiungo il numero percentuale alle etichette**
- **lbls <- paste(lbls, "% ", sep=" ") # aggiungo il simbolo % alle etichette**
- **pie(slices, labels = lbls, main="Grafico a torta dei partiti presenti all'EuroParlamento 2009")**

Grafici a torta

Grafico a torta dei partiti presenti all'EuroParlamento 2009

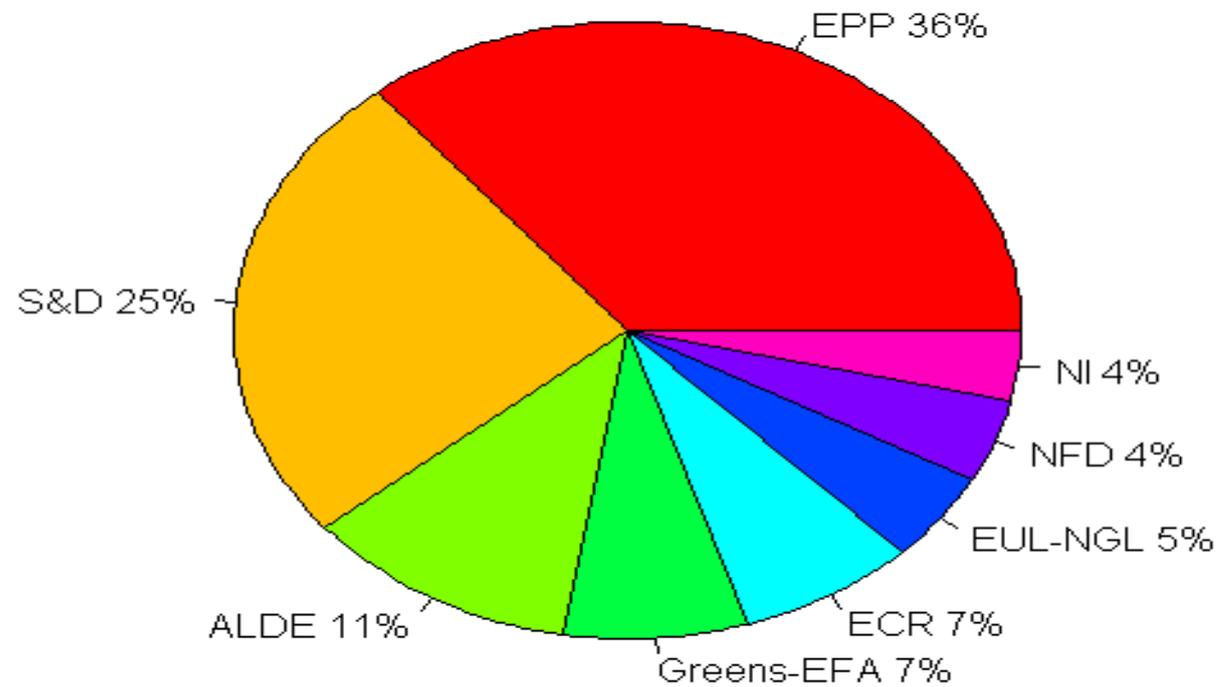


Grafici a torta

- Senza ulteriori specificazioni, ogni fetta della torta viene colorata con un colore leggermente diverso.
- Possiamo cambiare questa impostazione, ad esempio scegliendo un unico colore per ogni fetta (è sufficiente aggiungere, ad esempio, `col="red"`)
- Oppure possiamo decidere di colorare diversamente ogni fetta, con il seguente comando: `col=rainbow(length(lbls))`

Grafici a torta

Grafico a torta dei partiti presenti all'EuroParlamento 2009



Grafici a torta 3D

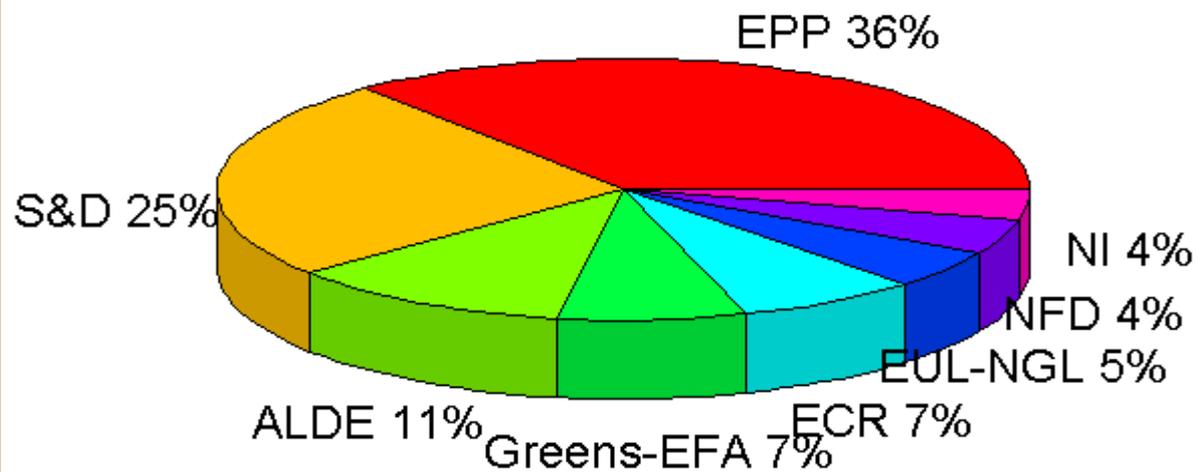
- Per creare un grafico a torta tridimensionale, dobbiamo caricare il package **plotrix**.

```
> library(plotrix)
```

```
> pie3D(slices, labels = lbls, main="Grafico a torta dei partiti  
presenti all'EuroParlamento 2009")
```

Grafici a torta 3D

Grafico a torta dei partiti presenti all'EuroParlamento 2009



Grafici a torta 3D

- Questo grafico può essere ampiamente personalizzato
- Possiamo ad esempio specificare l'altezza della torta (`height=`) o il suo raggio (`radius=`)
- l'angolo a partire dal quale cominciare a tagliare le fette espresso in radianti (`start=`),
- a quale angolatura presentare la torta (`theta=`)
- I colori (`col=`)
- la posizione delle etichette (`labelpos=`) e il loro colore (`labelcol=`)
- e possiamo scegliere di separare le fette con l'argomento `explode`

Esercizi

3. I seguenti dati rappresentano le dimensioni di 30 famiglie
5, 13, 9, 12, 7, 4, 8, 6, 6, 10, 7, 11, 10, 8, 15,
8, 6, 9, 12, 10, 7, 11, 10, 8, 12, 9, 7, 10, 7, 8
- a) Costruisci una tabella delle frequenze per questi dati
 - b) Rappresenta i dati con un grafico a bastoncini
 - c) Rappresenta i dati con un grafico poligonale
4. Indica se ciascuno dei seguenti insiemi di dati è simmetrico, approssimativamente simmetrico, o decisamente non simmetrico
- A: 6, 0, 2, 1, 8, 3, 5; B: 4, 0, 4, 0, 2, 1, 3, 2
C: 1, 1, 0, 1, 0, 3, 3, 2, 2, 2; D: 9, 9, 1, 2, 3, 9, 8, 4, 5

Esercizi

5. I seguenti numeri indicano la concentrazione di ozono nell'aria durante 25 giorni consecutivi

6.2, 9.1, 2.4, 3.6, 1.9, 1.7, 4.5, 4.2, 3.3, 5.1, 6.0, 1.8, 2.3, 4.9, 3.7, 3.8, 5.5, 6.4, 8.6, 9.3, 7.7, 5.4, 7.2, 4.9, 6.2

- Costruisci un istogramma delle frequenze usando la classe da 3 a 5
- Costruisci un istogramma delle frequenze usando la classe da 2 a 3
- Quale istogramma delle frequenze ti sembra più utile?

6. Il numero di decessi su strade lo scorso anno è il seguente:
Pedoni: 1699, Ciclisti: 280, Motociclisti:650, automobilisti: 1327
- Rappresenta questo insieme di dati in un grafico a torta assegnando il colore rosso ai Pedoni, ed il verde ai Motociclisti.