

```
library(rmf)#istallo il pacchetto rmf
dir("R")#mi restituisce il nome delle cartelle o file presenti all'interno della directory nominata
#Carico il file wvs
load("~/Documents/Informatica/R/ESERCIZIO DA FARE A CASA/homework ADES
2022-2023/WorldValueSurvey.rdata")
is(wvs)#verifico che abbia installato correttamente il file contenente i dati
```

## #ESERCITAZIONE

### #PUNTO 1

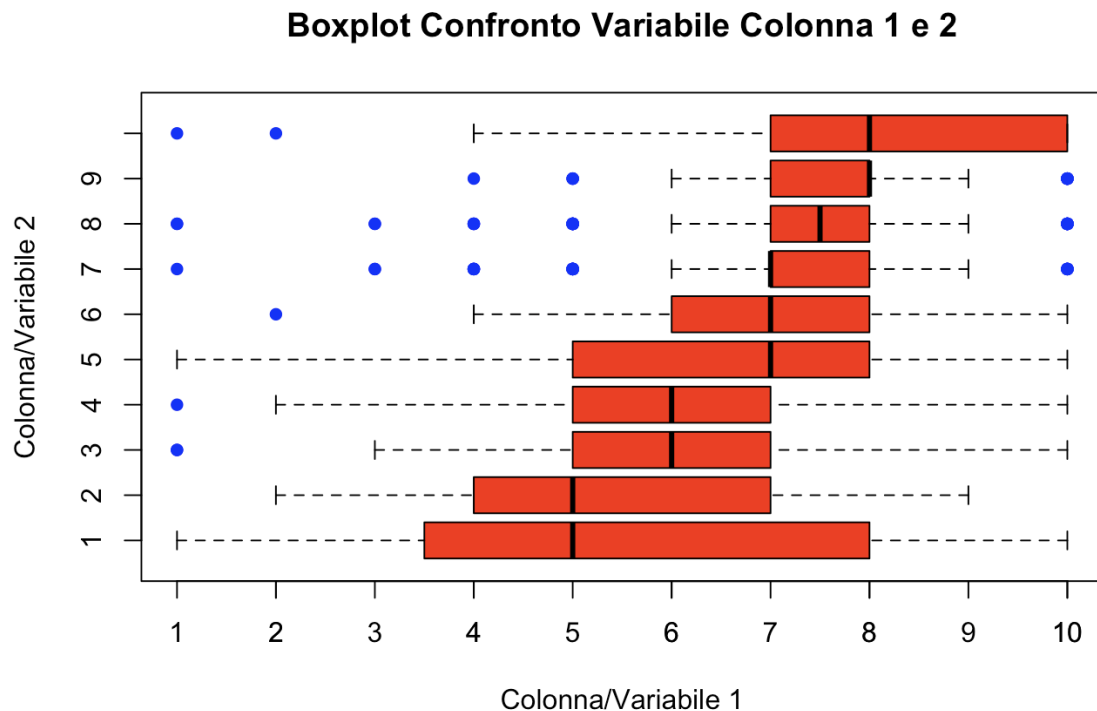
```
#Si può utilizzare anche la dicitura wvs[,1] per selezionare una determinata colonna o riga
#verifico il numero di righe che, con la prossima funzione, dovrà corrispondere alla somma
delle frequenze assolute più le osservazioni mancanti
length(wvs$v22)
#Calcolo la distribuzione di frequenza
#utilizzo la funzione "frequenze" fornita dal pacchetto rmf per calcolare la distribuzione di
frequenza della variabile la cui posizione nel data-frame wvs è riportata nella colonna (1)
frequenze(wvs$v22,mult = 1)
#utilizzo il parametro mult che mi permette di definire le posizioni della virgola e quindi
definire correttamente le frequenze relative (somma uguale a 1)
#dalla precedente funzione sono risultate 6 osservazioni mancanti che sommate alle 1006
calcolate portano alla lunghezza effettiva della colonna
#utilizzo il parametro na.rm per trascurare i dati mancanti (non numerici)
#Calcolo la media
M <- mean(wvs$v22, na.rm=TRUE)# media calcolata con il modulo base di R e l'associa ad
una variabile
M
#Calcolo la mediana
Me <- median(wvs$v22, na.rm=TRUE)# mediana calcolata con il modulo base di R e
l'associa ad una variabile
Me
#Calcolo la varianza
v <- var(wvs$v22, na.rm=TRUE)# varianza calcolata con il modulo base di R e l'associa ad
una variabile
v
#Calcolo la deviazione standard o scarto quadratico medio
ds <- sqrt(v)#sfrutto l'operatore matematico della radice quadrata
ds
#Utilizzo la funzione rbind per allineare orizzontalmente i valori richiesti e quelli calcolati
in modo da avere i dati organizzati
r<-rbind(c("Media","Mediana","Varianza","Deviazione Standard"), c(M,Me,v,ds),
deparse.level = 0)
r
```

## #PUNTO 2

```
#utilizzo la funzione tapply per applicare una funzione specifica ad un vettore o ad un fattore
#suddiviso in gruppi definiti da un altro vettore o fattore, applico quindi la funzione media
sulla prima colonna dipendente dalla seconda
t<-tapply(wvs$v22,wvs$v46,mean,na.rm=TRUE)
t
```

## #PUNTO 3

```
#Creo il boxplot utilizzando la tilde in modo da confrontare la distribuzione della prima
variabile/colonna dipendente dalla seconda variabile/colonna
boxplot(wvs$v22 ~ wvs$v46,wvs,range = 1.5, horizontal = TRUE, main = "Boxplot Confronto
Variabile Colonna 1 e 2", xlab = "Colonna/Variabile 1", ylab = "Colonna/Variabile 2",col=
"red", outcol = "blue", pch = 16)
axis(side = 1, at = 1:10, labels = 1:10)#numero da 1 a 10 l'asse x
```



## #PUNTO 4

```
#Calcolo la covarianza utilizzando la formula base di R
co<-cov(wvs$v47,wvs$v68, use="complete.obs")
co
#Calcolo il coefficiente di correlazione lineare tra due variabili continue quindi il coefficiente
di Pearson utilizzando la funzione base di R
cr<-cor(wvs$v47,wvs$v68,use = "complete.obs",method = "pearson")
cr
```

```

#PUNTO 5
#Risolvo l'esercizio attraverso l'utilizzzo del metodo di Monte Carlo
#definisco il numero di facce del dado che corrispondono al numero compreso nella colonna
5
dado <- 20
#definisco il numero di ripetizioni dell'esperimento
N <- 100000
#creo un vettore nullo dove poter scrivere il valore più alto ottenuto ad ogni esperimento
prove <- numeric(N)
#utilizzo il ciclo for per ripetere l'operazione tante volte quante N
for(j in 1:N)
{
  #calcolo il valore massimo dell'esperimento corrente lanciando il dado un numero di volte
  pari alla colonna 6
  x <- max(sample(dado,size = 3),replace = TRUE)
  #scrivo nel vettore vuoto, i valori massimi ottenuti
  prove[j] <- x
}
#rappresento la tabella di distribuzione contenente anche la frequenza cumulativa
frequenze(prove,cumul = TRUE, mult = 1)
#definisco una variabile alla quale si assegna il vincolo (condizione): vettore prove maggiore
di colonna 7
ok <- prove >= 12
#assegno ad una variabile (out), solo i valori del vettore dei massimi (prove) che
rispecchiano la condizione della riga precedente
out <- prove[ok]
#calcolo i casi favorevoli che corrispondono al numero di elementi della variabile out
Cf<-length(out)
#utizzo una if e in caso Cf fosse maggiore di 0 allora calcolo la probabilità, nei casi rimanenti
vengo informato del risultato specifico
if (Cf > 0) {
  probabilita <- Cf/N*100
  cat("La probabilità che il massimo raggiunga o superi il valore riportato nella colonna 7 è:",
  probabilita, "%")
} else if(Cf < 0){
  cat("Probabilità < 0. La probabilità non è calcolabile.")
} else {
  cat("Probabilità = 0. La probabilità non è calcolabile.")
}

```

#### #PUNTO 6

#definisco la probabilità di insolvenza di ciascun credito

pi <- 0.04

#definisco il numero complessivo di crediti

nc <- 262

#definisco il numero di insolvenze della da superare cioè quello della colonna 10

nic <- 9

#calcolo il numero di insolvenze attese

VA <- pi\*nc

VA

#calcolo le probabilità di non superare il valore atteso

pVA <- pbinom(VA,nc,pi)

pVA

#calcolo le probabilità di non superare il valore della colonna 10

pnic <- pbinom(nic,nc,pi)

pnic

#organizzo i dati

table <- cbind(Categoria = c("Valore Atteso", "Valore Da Superare"), Valore = c(VA,nic),  
ProbabilitàMinUguale = c(pVA,pnic), ProbabilitàMaggiore = c(1-pVA, 1-pnic))

table

#### #PUNTO 7

#definisco il valore di mu

muES <- 9.9

#definisco il valore di sigma

sigmaES <- 11.6

#definisco il valore massimo

valConf <- 31.5

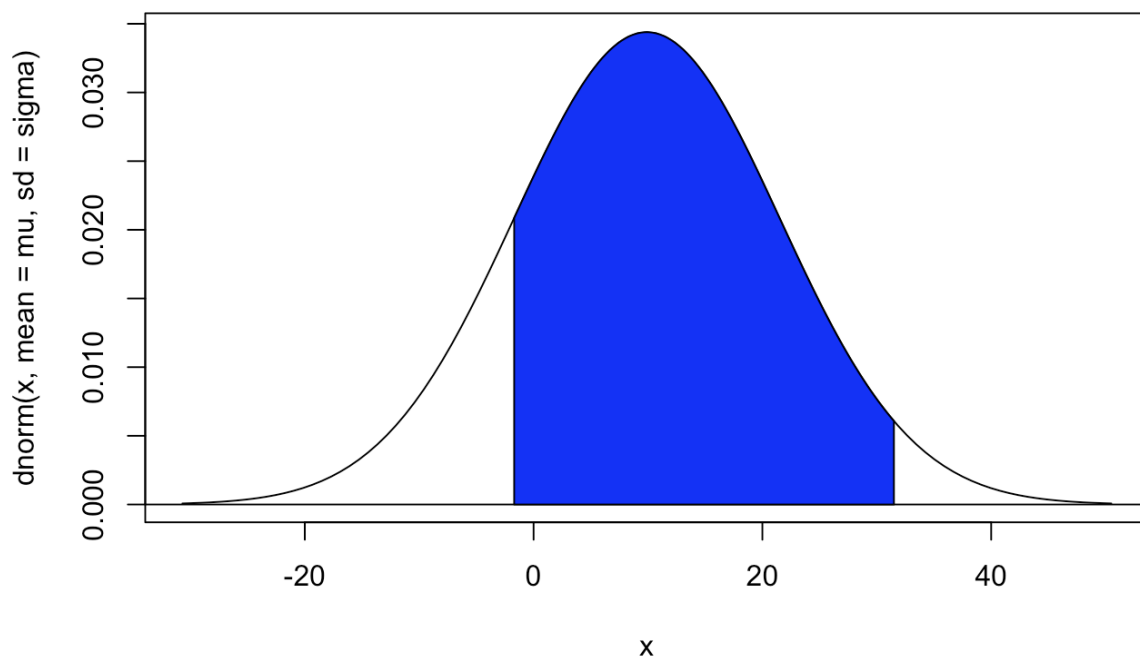
#calcolo le probabilità cumulative della Normale utilizzando la funzione ProbNorm del  
pacchetto rmf che mi permette anche la rappresentazione grafica

ProbNorm(da = muES - sigmaES, a = valConf, mu = muES, sigma = sigmaES, color  
="blue")

#calcolo le probabilità cumulative della Normale utilizzando la funzione pnorm di R

prob <- pnorm(valConf, mean = muES, sd = sigmaES) - pnorm(muES - sigmaES, mean =  
muES, sd = sigmaES)

prob



```
#PUNTO 8
#Definisco il valore minimo
Min <- -3
#Definisco il valore massimo
Max <- 5
#Inizializzo il seme
set.seed(236704)
#Creo il campione casuale e dipendente dal seme
x<- sample(Min:Max,size=300,replace=TRUE)
#Resetto il campionamento
set.seed(NULL)
#Calcolo la distribuzione di frequenze del campione generato
f <- frequenze(x)
f
```

```
#Punto A
#Calcolo la mediana utilizzando la distribuzione di frequenza
MeF <- (min(x) + max(x)/2)
MeF
#Calcolo la mediana utilizzando i dati grezzi
MeG <-median(x,na.rm = TRUE)
MeG
#Creo una tabella per semplificare la visione dei dati
```

```
t <- cbind(c("Mediana di Distribuzione di Frequenza","Mediana di Dati Grezzi"), c(MeF,MeG))
t
```

```
#Punto B
#Calcolo la media degli scarti in valore assoluto dalla mediana
SMe <- mean(abs(x-MeG))
SMe
#Calcolo la media degli scarti in valore assoluto utilizzando un altro valore
A <- mean(abs(x-(-7)))
A
SMe < A#la teoria risulta verificata
#Calcolo nuovamente la media degli scarti in valore assoluto utilizzando un ulteriore valore
B <- mean(abs(x-20))
B
SMe < B#la teoria risulta verificata nuovamente
```

```
#PUNTO 9
#Carico il file contenente la tabella dei dati covid dopo aver cambiato la fascia di età errata
della riga due del file excel
library(readxl)
dati_covid <- read_excel("Documents/Informatica/R/ESERCIZIO DA FARE A
CASA/homework ADES 2022-2023/dati covid.xlsx")
#Apro il foglio contenente il data frame originale
View(dati_covid)
#Creo un data frame senza la prima e ultima riga in modo da poter sfruttare in maniera
efficiente l'operatore $
dati_covid2 <- subset(dati_covid, !(rownames(dati_covid) == 1 | rownames(dati_covid) ==
12))
#Apro un nuovo foglio contenente il nuovo data frame
View(dati_covid2)
#Organizzo i dati di sintesi grafica
par(mfrow=c(2,3))
```

```
#RICHIESTA 1
# Creo il vettore delle fasce d'età
Fde <- c(0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100)
```

```
#MASCHI
```

```
# Utilizzo la funzione scipen per modificare le opzioni e scrivere i valori dell'asse y in valori
numerici e non in notazione scientifica
options(scipen = 999)
# Creo l'istogramma di frequenza
Freq <- c(rep(5, dati_covid2[1, 3]), rep(15, dati_covid2[2, 3]), rep(25, dati_covid2[3, 3]),
rep(35, dati_covid2[4, 3]), rep(45, dati_covid2[5, 3]), rep(55, dati_covid2[6, 3]), rep(65,
dati_covid2[7, 3]), rep(75, dati_covid2[8, 3]), rep(85, dati_covid2[9, 3]), rep(95,
dati_covid2[10, 3]))
# Creo l'istogramma con gli assi personalizzati
hist(Freq, breaks = Fde, main = "Età al Contagio Maschile", xlab = "Classe di Età", ylab =
"Numero di Casi", col = "blue", axes = FALSE)
# Aggiungo gli assi personalizzati con le etichette complete
axis(1, at = Fde, labels = Fde)
axis(2)
# Aggiungo la griglia di riferimento
grid()
```

#### #FEMMINE

```
# Utilizzo la funzione scipen per modificare le opzioni e scrivere i valori dell'asse y in valori
numerici e non in notazione scientifica
options(scipen = 999)
# Creo l'istogramma di frequenza
Freq <- c(rep(5, dati_covid2[1, 8]), rep(15, dati_covid2[2, 8]), rep(25, dati_covid2[3, 8]),
rep(35, dati_covid2[4, 8]), rep(45, dati_covid2[5, 8]), rep(55, dati_covid2[6, 8]), rep(65,
dati_covid2[7, 8]), rep(75, dati_covid2[8, 8]), rep(85, dati_covid2[9, 8]), rep(95,
dati_covid2[10, 8]))
# Creo l'istogramma con gli assi personalizzati
hist(Freq, breaks = Fde, main = "Età al Contagio Femminile", xlab = "Classe di Età", ylab =
"Numero di Casi", col = "red", axes = FALSE)
# Aggiungo gli assi personalizzati con le etichette complete
axis(1, at = Fde, labels = Fde)
axis(2)
# Aggiungo la griglia di riferimento
grid()
```

#### #RICHIESTA 2

#### #MASCHI

```
#Calcolo la MEDIANA della variabile età al contagio
#Estremo inferiore della classe mediana
lm <- 40
#Frequenza relativa cumulata fino alla classe mediana
```

```

Fm <- sum(as.numeric(dati_covid2$...3[1:6]))
#Frequenza relativa cumulata fino alla classe precedente a quella mediana
Fm_1 <- sum(as.numeric(dati_covid2$...3[1:4]))
#Ampiezza della classe mediana
DeltaM <- 10
#Numero totale di casi maschili
TotCasi <- as.numeric(dati_covid[12,3])
#Formula di risoluzione
MeM <- lm+((0.5-Fm_1/TotCasi)/(Fm/TotCasi-Fm_1/TotCasi))*DeltaM
message1 <- paste("L'età mediana della totalità dei casi maschili è:", MeM)
print(message1)

```

```

#Calcolo la MEDIA della variabile età al contagio utilizzando la formula della media
ponderata
#Creo il vettore delle medie delle fasce d'età
Mcde <- c(5,15,25,35,45,55,65,75,85,95)
#Definisco il vettore età
xci <- Mcde
#Seleziono la colonna del numero dei casi e faccio in modo che sia di tipo numerico
ni <- as.numeric(dati_covid2$...3)
#calcolo la media secondo la formula della media ponderata
MM <- (sum(xci*ni))/sum(ni)
message2 <- paste("L'età media della totalità dei casi maschili è:", MM)
print(message2)

```

#Tra le due forme di tendenza centrale precedentemente calcolate scelgo la mediana.  
 #Questa scelta è dovuta dal fatto che, in questo caso specifico, la mediana esprima meglio la tendenza centrale dei casi in confronto alla media.  
 #Esprime meglio la tendenza centrale perchè la media è influenzata dagli outlier non permettendo così una misura di centralità adeguata che rispecchi appunto la centralità dei dati stessi

```

#Calcolo lo SCARTO QUADRATICO MEDIO della variabile età al contagio utilizzando la
formula apposita per le distribuzioni di frequenze
#Creo il vettore delle medie delle fasce d'età
Mcde <- c(5,15,25,35,45,55,65,75,85,95)
#Definisco il vettore età
xci <- Mcde
#Seleziono la colonna del numero dei casi e faccio in modo che sia di tipo numerico
ni <- as.numeric(dati_covid2$...3)
#calcolo la media secondo la formula della media ponderata
DsM <- sqrt((sum(((xci-MM)^2)*ni))/sum(ni))
message3 <- paste("Lo scarto quadratico medio del numero totale di casi maschili
rapportato all'età è:", DsM)
print(message3)

```



#FEMMINE

```
#Calcolo la MEDIANA della variabile età al contagio
#Estremo inferiore della classe mediana
Im <- 40
#Frequenza relativa cumulata fino alla classe mediana
Fm <- sum(as.numeric(dati_covid2$...8[1:6]))
#Frequenza relativa cumulata fino alla classe precedente a quella mediana
Fm_1 <- sum(as.numeric(dati_covid2$...8[1:4]))
#Ampiezza della classe mediana
DeltaM <- 10
#Numero totale di casi maschili
TotCasi <- as.numeric(dati_covid[12,8])
#Formula di risoluzione
MeF <- Im+((0.5-Fm_1/TotCasi)/(Fm/TotCasi-Fm_1/TotCasi))*DeltaM
message4 <- paste("L'età mediana della totalità dei casi femminili è:", MeF)
print(message4)
```

```
#Calcolo la MEDIA della variabile età al contagio utilizzando la formula della media
ponderata
#Creo il vettore delle medie delle fasce d'età
Mcde <- c(5,15,25,35,45,55,65,75,85,95)
#Definisco il vettore età
xci <- Mcde
#Selezione la colonna del numero dei casi e faccio in modo che sia di tipo numerico
ni <- as.numeric(dati_covid2$...8)
#calcolo la media secondo la formula della media ponderata
MF <- (sum(xci*ni))/sum(ni)
message5 <- paste("L'età media della totalità dei casi femminili è:", MF)
print(message5)
```

```
#Calcolo lo SCARTO QUADRATICO MEDIO della variabile età al contagio utilizzando la
formula apposita per le distribuzioni di frequenze
#Creo il vettore delle medie delle fasce d'età
Mcde <- c(5,15,25,35,45,55,65,75,85,95)
#Definisco il vettore età
xci <- Mcde
#Selezione la colonna del numero dei casi e faccio in modo che sia di tipo numerico
ni <- as.numeric(dati_covid2$...8)
#calcolo la media secondo la formula della media ponderata
DsF <- sqrt((sum(((xci-MF)^2)*ni))/sum(ni))
message6 <- paste("Lo scarto quadratico medio del numero totale di casi femminili
rapportato all'età è:", DsF)
print(message6)
```

### #RICHIESTA 3

#Sommo i casi maschili a quelli femminili

```
s <- as.numeric(dati_covid2$...3) + as.numeric(dati_covid2$...8)
```

#Sommo i casi totali maschili a quelli femminili

```
t <- as.numeric(dati_covid[12,3]) + as.numeric(dati_covid[12,8])
```

#Calcolo l'ISTOGRAMMA DI FREQUENZA della variabile età al contagio dei casi totali

# Creo il vettore delle fasce d'età

```
Fde <- c(0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100)
```

# Utilizzo la funzione scipen per modificare le opzioni e scrivere i valori dell'asse y in valori numerici e non in notazione scientifica

```
options(scipen = 999)
```

# Creo l'istogramma di frequenza

```
Freq <- c(rep(5, s[1]), rep(15, s[2]), rep(25, s[3]), rep(35, s[4]), rep(45, s[5]), rep(55, s[6]),  
rep(65, s[7]), rep(75, s[8]), rep(85, s[9]), rep(95, s[10]))
```

# Creo l'istogramma con gli assi personalizzati

```
hist(Freq, breaks = Fde, main = "Età al Contagio Totale", xlab = "Classe di Età", ylab =  
"Numero di Casi Totali", col = "green", axes = FALSE)
```

# Aggiungo gli assi personalizzati con le etichette complete

```
axis(1, at = Fde, labels = Fde)
```

```
axis(2)
```

# Aggiungo la griglia di riferimento

```
grid()
```

#Calcolo la MEDIANA della variabile età al contagio dei casi totali

#Estremo inferiore della classe mediana

```
lm <- 40
```

#Frequenza relativa cumulata fino alla classe mediana

```
Fm <- sum(as.numeric(s[1:6]))
```

#Frequenza relativa cumulata fino alla classe precedente a quella mediana

```
Fm_1 <- sum(as.numeric(s[1:4]))
```

#Ampiezza della classe mediana

```
DeltaM <- 10
```

#Numero totale di casi maschili

```
TotCasi <- t
```

#Formula di risoluzione

```
MeT <- lm + ((0.5 - Fm_1 / TotCasi) / (Fm / TotCasi - Fm_1 / TotCasi)) * DeltaM
```

```
message7 <- paste("L'età mediana della totalità dei casi è:", MeT)
```

```
print(message7)
```

#Calcolo la MEDIA della variabile età al contagio dei casi totali utilizzando la formula della media ponderata

#Creo il vettore delle medie delle fasce d'età

```
Mcde <- c(5,15,25,35,45,55,65,75,85,95)
```

#Definisco il vettore età

```
xcj <- Mcde
```

#Seleziono la colonna del numero dei casi e faccio in modo che sia di tipo numerico

```

ni <- s
#calcolo la media secondo la formula della media ponderata
MT <- (sum(xci*ni))/sum(ni)
message8 <- paste("L'età media della totalità dei casi è:", MT)
print(message8)

#Calcolo lo SCARTO QUADRATICO MEDIO della variabile età al contagio dei casi totali
utilizzando la formula apposita per le distribuzioni di frequenze
#Creo il vettore delle medie delle fasce d'età
Mcde <- c(5,15,25,35,45,55,65,75,85,95)
#Definisco il vettore età
xci <- Mcde
#Selezione la colonna del numero dei casi e faccio in modo che sia di tipo numerico
ni <- s
#calcolo la media secondo la formula della media ponderata
DsT <- sqrt((sum(((xci-MT)^2)*ni))/sum(ni))
message9 <- paste("Lo scarto quadratico medio del numero totale di casi rapportato all'età
è:", DsT)
print(message9)

```

#### #RICHIESTA 4

```

#La mediana risulta maggiore perchè nel bollettino sono state contate anche fasce d'età
maggiori di 100.
#Un maggior numero di fasce d'età,utilizzando la formula apposita della mediana, porta ad
un risultato maggiore

```

#### #RICHIESTA 5

##### #Punto I

##### #MASCHI

```

# Creo il vettore delle fasce d'età
Fde <- c(0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100)

#Calcolo l'ISTOGRAMMA DI FREQUENZA della variabile età al decesso maschile
# Utilizzo la funzione scipen per modificare le opzioni e scrivere i valori dell'asse y in valori
numerici e non in notazione scientifica
options(scipen = 999)
# Creo l'istogramma di frequenza
Freq <- c(rep(5, dati_covid2[1, 2]), rep(15, dati_covid2[2, 2]), rep(25, dati_covid2[3, 2]),
rep(35, dati_covid2[4, 2]), rep(45, dati_covid2[5, 2]), rep(55, dati_covid2[6, 2]), rep(65,
dati_covid2[7, 2]), rep(75, dati_covid2[8, 2]), rep(85, dati_covid2[9, 2]), rep(95,
dati_covid2[10, 2]))

```

```
# Creo l'istogramma con gli assi personalizzati
hist(Freq, breaks = Fde, main = "Età al Decesso Maschile", xlab = "Classe di Età", ylab =
"Numero di Decessi", col = "blue", axes = FALSE)
# Aggiungo gli assi personalizzati con le etichette complete
axis(1, at = Fde, labels = Fde)
axis(2)
# Aggiungo la griglia di riferimento
grid()
```

```
#FEMMINE
```

```
#Calcolo l'ISTOGRAMMA DI FREQUENZA della variabile età al decesso femminile
# Utilizzo la funzione scipen per modificare le opzioni e scrivere i valori dell'asse y in valori
numerici e non in notazione scientifica
options(scipen = 999)
# Creo l'istogramma di frequenza
Freq <- c(rep(5, dati_covid2[1, 7]), rep(15, dati_covid2[2, 7]), rep(25, dati_covid2[3, 7]),
rep(35, dati_covid2[4, 7]), rep(45, dati_covid2[5, 7]), rep(55, dati_covid2[6, 7]), rep(65,
dati_covid2[7, 7]), rep(75, dati_covid2[8, 7]), rep(85, dati_covid2[9, 7]), rep(95,
dati_covid2[10, 7]))
# Creo l'istogramma con gli assi personalizzati
hist(Freq, breaks = Fde, main = "Età al Decesso Femminile", xlab = "Classe di Età", ylab =
"Numero di Decessi", col = "red", axes = FALSE)
# Aggiungo gli assi personalizzati con le etichette complete
axis(1, at = Fde, labels = Fde)
axis(2)
# Aggiungo la griglia di riferimento
grid()
```

```
#Punto II
```

```
#MASCHI
```

```
#Calcolo la MEDIANA della variabile età al decesso
#Estremo inferiore della classe mediana
Im <- 40
#Frequenza relativa cumulata fino alla classe mediana
Fm <- sum(as.numeric(dati_covid2$...2[1:6]))
#Frequenza relativa cumulata fino alla classe precedente a quella mediana
Fm_1 <- sum(as.numeric(dati_covid2$...2[1:4]))
#Ampiezza della classe mediana
DeltaM <- 10
#Numero totale di decessi maschili
TotCasi <- as.numeric(dati_covid[12,2])
#Formula di risoluzione
MeMD <- Im+((0.5-Fm_1/TotCasi)/(Fm/TotCasi-Fm_1/TotCasi))*DeltaM
message10 <- paste("L'età mediana della totalità dei decessi maschili è:", MeMD)
print(message10)
```

```

#Calcolo la MEDIA della variabile età al decesso utilizzando la formula della media
ponderata
#Creo il vettore delle medie delle fasce d'età
Mcde <- c(5,15,25,35,45,55,65,75,85,95)
#Definisco il vettore età
xci <- Mcde
#Seleziono la colonna del numero dei casi e faccio in modo che sia di tipo numerico
ni <- as.numeric(dati_covid2$...2)
#calcolo la media secondo la formula della media ponderata
MMD <- (sum(xci*ni))/sum(ni)
message11 <- paste("L'età media della totalità dei decessi maschili è:", MMD)
print(message11)

```

```

#Calcolo lo SCARTO QUADRATICO MEDIO della variabile età al decesso utilizzando la
formula aposita per le distribuzioni di frequenze
#Creo il vettore delle medie delle fasce d'età
Mcde <- c(5,15,25,35,45,55,65,75,85,95)
#Definisco il vettore età
xci <- Mcde
#Seleziono la colonna del numero dei casi e faccio in modo che sia di tipo numerico
ni <- as.numeric(dati_covid2$...2)
#calcolo la media secondo la formula della media ponderata
DsMD <- sqrt((sum(((xci-MMD)^2)*ni))/sum(ni))
message12 <- paste("Lo scarto quadratico medio del numero totale di decessi maschili
rapportato all'età è:", DsMD)
print(message12)

```

```

#FEMMINE
#Calcolo la MEDIANA della variabile età al decesso
#Estremo inferiore della classe mediana
Im <- 40
#Frequenza relativa cumulata fino alla classe mediana
Fm <- sum(as.numeric(dati_covid2$...7[1:6]))
#Frequenza relativa cumulata fino alla classe precedente a quella mediana
Fm_1 <- sum(as.numeric(dati_covid2$...7[1:4]))
#Ampiezza della classe mediana
DeltaM <- 10
#Numero totale di casi maschili
TotCasi <- as.numeric(dati_covid[12,7])
#Formula di risoluzione
MeFD <- Im+((0.5-Fm_1/TotCasi)/(Fm/TotCasi-Fm_1/TotCasi))*DeltaM
message13 <- paste("L'età mediana della totalità dei decessi femminili è:", MeFD)
print(message13)

```

```

#Calcolo la MEDIA della variabile età al decesso utilizzando la formula della media
ponderata
#Creo il vettore delle medie delle fasce d'età
Mcde <- c(5,15,25,35,45,55,65,75,85,95)

```

```

#Definisco il vettore età
xci <- Mcde
#Seleziono la colonna del numero dei casi e faccio in modo che sia di tipo numerico
ni <- as.numeric(dati_covid2$...7)
#calcolo la media secondo la formula della media ponderata
MFD <- (sum(xci*ni))/sum(ni)
message14 <- paste("L'età media della totalità dei decessi femminili è:", MFD)
print(message14)

#Calcolo lo SCARTO QUADRATICO MEDIO della variabile età al decesso utilizzando la
formula apposita per le distribuzioni di frequenze
#Creo il vettore delle medie delle fasce d'età
Mcde <- c(5,15,25,35,45,55,65,75,85,95)
#Definisco il vettore età
xci <- Mcde
#Seleziono la colonna del numero dei casi e faccio in modo che sia di tipo numerico
ni <- as.numeric(dati_covid2$...7)
#calcolo la media secondo la formula della media ponderata
DsFD <- sqrt((sum(((xci-MFD)^2)*ni))/sum(ni))
message15 <- paste("Lo scarto quadratico medio del numero totale di decessi femminili
rapportato all'età è:", DsFD)
print(message15)

```

```

#Punto III
#Sommo i decessi maschili a quelli femminili
ss <- as.numeric(dati_covid2$...2) + as.numeric(dati_covid2$...7)
#Sommo i decessi totali maschili a quelli femminili
tt <- as.numeric(dati_covid[12,2]) + as.numeric(dati_covid[12,7])

#Calcolo l'ISTOGRAMMA DI FREQUENZA della variabile età al decesso dei casi totali
# Creo il vettore delle fasce d'età
Fde <- c(0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100)
# Utilizzo la funzione scipen per modificare le opzioni e scrivere i valori dell'asse y in valori
numerici e non in notazione scientifica
options(scipen = 999)
# Creo l'istogramma di frequenza
Freq <- c(rep(5, ss[1]), rep(15, ss[2]), rep(25, ss[3]), rep(35, ss[4]), rep(45, ss[5]), rep(55,
ss[6]), rep(65, ss[7]), rep(75, ss[8]), rep(85, ss[9]), rep(95, ss[10]))
# Creo l'istogramma con gli assi personalizzati
hist(Freq, breaks = Fde, main = "Età al Decesso Totale", xlab = "Classe di Età", ylab =
"Numero di Decessi Totali", col = "green", axes = FALSE)
# Aggiungo gli assi personalizzati con le etichette complete
axis(1, at = Fde, labels = Fde)
axis(2)
# Aggiungo la griglia di riferimento
grid()

```

```

#Calcolo la MEDIANA della variabile età al decesso dei decessi totali
#Estremo inferiore della classe mediana
Im <- 40
#Frequenza relativa cumulata fino alla classe mediana
Fm <- sum(as.numeric(ss[1:6]))
#Frequenza relativa cumulata fino alla classe precedente a quella mediana
Fm_1 <- sum(as.numeric(ss[1:4]))
#Ampiezza della classe mediana
DeltaM <- 10
#Numero totale di casi maschili
TotCasi <- tt
#Formula di risoluzione
MeTD <- Im+((0.5-Fm_1/TotCasi)/(Fm/TotCasi-Fm_1/TotCasi))*DeltaM
message16 <- paste("L'età mediana della totalità dei decessi è:", MeTD)
print(message16)

```

```

#Calcolo la MEDIA della variabile età al decesso dei casi totali utilizzando la formula della
media ponderata
#Creo il vettore delle medie delle fasce d'età
Mcde <- c(5,15,25,35,45,55,65,75,85,95)
#Definisco il vettore età
xci <- Mcde
#Seleziono la colonna del numero dei casi e faccio in modo che sia di tipo numerico
ni <- ss
#calcolo la media secondo la formula della media ponderata
MTD <- (sum(xci*ni))/sum(ni)
message17 <- paste("L'età media della totalità dei decessi è:", MTD)
print(message17)

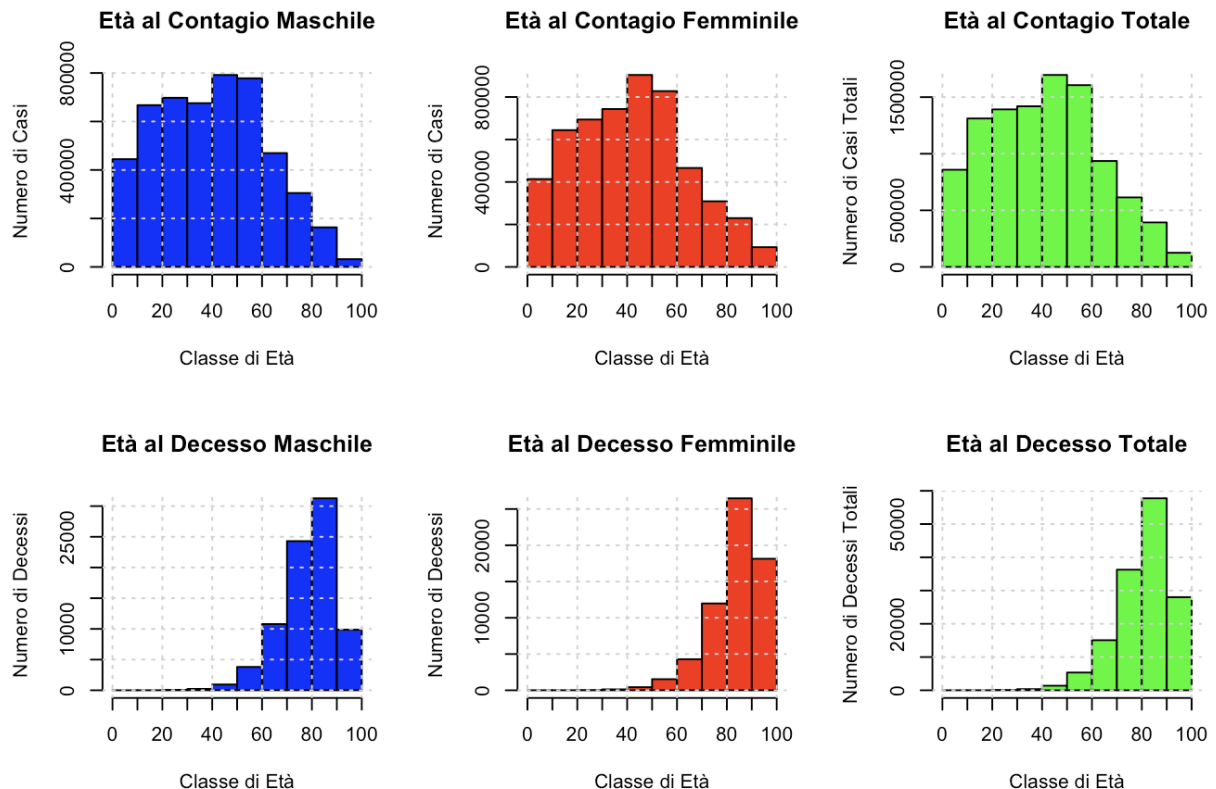
```

```

#Calcolo lo SCARTO QUADRATICO MEDIO della variabile età al contagio dei casi totali
utilizzando la formula apposita per le distribuzioni di frequenze
#Creo il vettore delle medie delle fasce d'età
Mcde <- c(5,15,25,35,45,55,65,75,85,95)
#Definisco il vettore età
xci <- Mcde
#Seleziono la colonna del numero dei casi e faccio in modo che sia di tipo numerico
ni <- ss
#calcolo la media secondo la formula della media ponderata
DsTD <- sqrt((sum(((xci-MTD)^2)*ni))/sum(ni))
message18 <- paste("Lo scarto quadratico medio del numero totale di decessi rapportato
all'età è:", DsTD)
print(message18)

```

#Torno al default per la costruzione di grafici dopo aver organizzato i dati di sintesi grafica  
`par(mfrow=c(1,1))`



#Per analizzare, successivamente, le variabilità dei differenti casi faccio affidamento al coefficiente di variabilità(CV) calcolato in %.

$$CVM = DsM/MM*100$$

$$CVF = DsF/MF*100$$

$$CVT = DsT/MT*100$$

$$CVMD = DsMD/MMD*100$$

$$CVFD = DsFD/MFD*100$$

$$CVTD = DsTD/MTD*100$$

#Organizzo i dati di sintesi numerica

#Organizzo i dati delle Misure di Centralità

```
dfMdC <- data.frame(
  M.CENTRALITÀ = c("Età Al Contagio Maschile", "Età Al Contagio Femminile", "Età Al Contagio Totale", "Età Al Decesso Maschile", "Età Al Decesso Femminile", "Età Al Decesso Totale"),
  M = c(MM, MF, MT, MMD, MFD, MTD),
  Me = c(MeM, MeF, MeT, MeMD, MeFD, MeTD)
)
dfMdC
View(dfMdC)
```



#Organizzo i dati delle Misure di Variabilità

```
dfMdV <- data.frame(  
  M.VARIABILITÀ = c("Età Al Contagio Maschile","Età Al Contagio Femminile","Età Al Contagio Totale",  
    "Età Al Decesso Maschile","Età Al Decesso Femminile","Età Al Decesso Totale"),  
  DS = c(DsM,DsF,DsT,DsMD,DsFD,DsTD),  
  CV = c(CVM,CVF,CVT,CVMD,CVFD,CVTD)  
)  
dfMdV  
View(dfMdV)
```

```
> #Organizzo i dati delle Misure di Centralità  
> dfMdC <- data.frame(  
+   M.CENTRALITÀ = c("Età Al Contagio Maschile","Età Al Contagio Femminile","Età Al Contagio Totale",  
+     "Età Al Decesso Maschile","Età Al Decesso Femminile","Età Al Decesso Totale"),  
+   M = c(MM,MF,MT,MMD,MFD,MTD),  
+   Me = c(MeM,MeF,MeT,MeMD,MeFD,MeTD)  
+ )  
> dfMdC
```

	M.CENTRALITÀ	M	Me
1	Età Al Contagio Maschile	40.21378	40.17502
2	Età Al Contagio Femminile	41.90715	40.96566
3	Età Al Contagio Totale	41.08494	40.58967
4	Età Al Decesso Maschile	78.49822	124.81918
5	Età Al Decesso Femminile	83.45690	199.51884
6	Età Al Decesso Totale	80.66384	146.65104

```
> View(dfMdC)
```

```
> #Organizzo i dati delle Misure di Variabilità  
> dfMdV <- data.frame(  
+   M.VARIABILITÀ = c("Età Al Contagio Maschile","Età Al Contagio Femminile","Età Al Contagio Totale",  
+     "Età Al Decesso Maschile","Età Al Decesso Femminile","Età Al Decesso Totale"),  
+   DS = c(DsM,DsF,DsT,DsMD,DsFD,DsTD),  
+   CV = c(CVM,CVF,CVT,CVMD,CVFD,CVTD)  
+ )  
> dfMdV
```

	M.VARIABILITÀ	DS	CV
1	Età Al Contagio Maschile	21.96736	54.62645
2	Età Al Contagio Femminile	22.47853	53.63889
3	Età Al Contagio Totale	22.24791	54.15100
4	Età Al Decesso Maschile	11.16397	14.22194
5	Età Al Decesso Femminile	10.69639	12.81666
6	Età Al Decesso Totale	11.23472	13.92782

```
> View(dfMdV)
```

#COMMENTO FINALE

#Analizzando le Misure di Centralità e partendo dal presupposto che in caso di simmetria, Media e Mediana coincidono;

#Si può notare dai grafici e dai calcoli che, nel caso di una distribuzione di variabili continue, maggiore è l'asimmetria destra (positiva) maggiore sarà la media in confronto alla mediana ottenendo così maggiore precisione di calcolo della mediana. E' dimostrato dai casi di contagio.

#Al contrario, nel caso dei decessi, vi è maggiore asimmetria sinistra (negativa) e ciò porta a una perdita di significato della misura mediana e una maggiore importanza del valore medio e in questo caso si avrà una mediana maggiore della media.

#Per le Misure di Variabilità si può notare che vi è una maggiore variabilità all'età al contagio in confronto all'età al decesso.

#Questo perché il contagio caratterizza una fascia d'età più ampia di quella dei decessi che invece caratterizza principalmente le fasce d'età più elevate.