class01intro.R

andre

Fri Jul 15 11:16:25 2016

```r
rm(list=ls())
cat("\014")
```

```r
age <- c(23, 35, 31, 49, 77, 42, 50, 30, 12)
salaryInThousands <- c(33, 105, 153, 0, 74, 99, 32, 188, 12)

mean(age)
```

```
## [1] 38.77778
```

```r
sd(age)
```

```
## [1] 18.7735
```
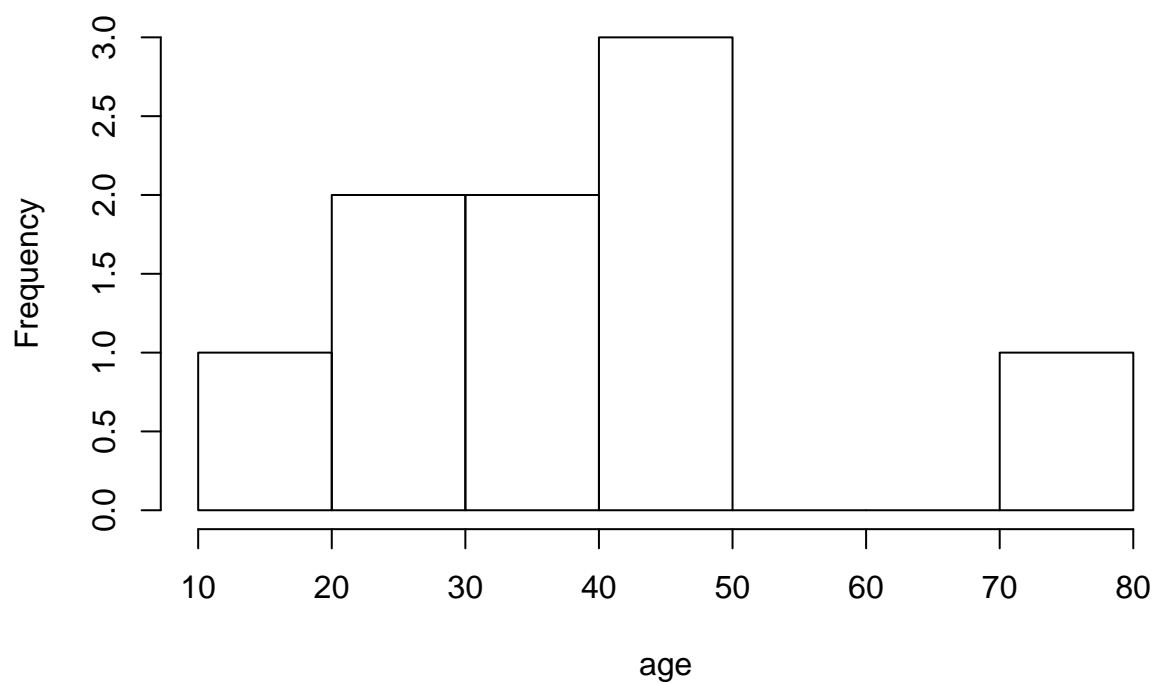
```r
var(salaryInThousands)
```

```
## [1] 4186
```

```r
cor(age,salaryInThousands)
```
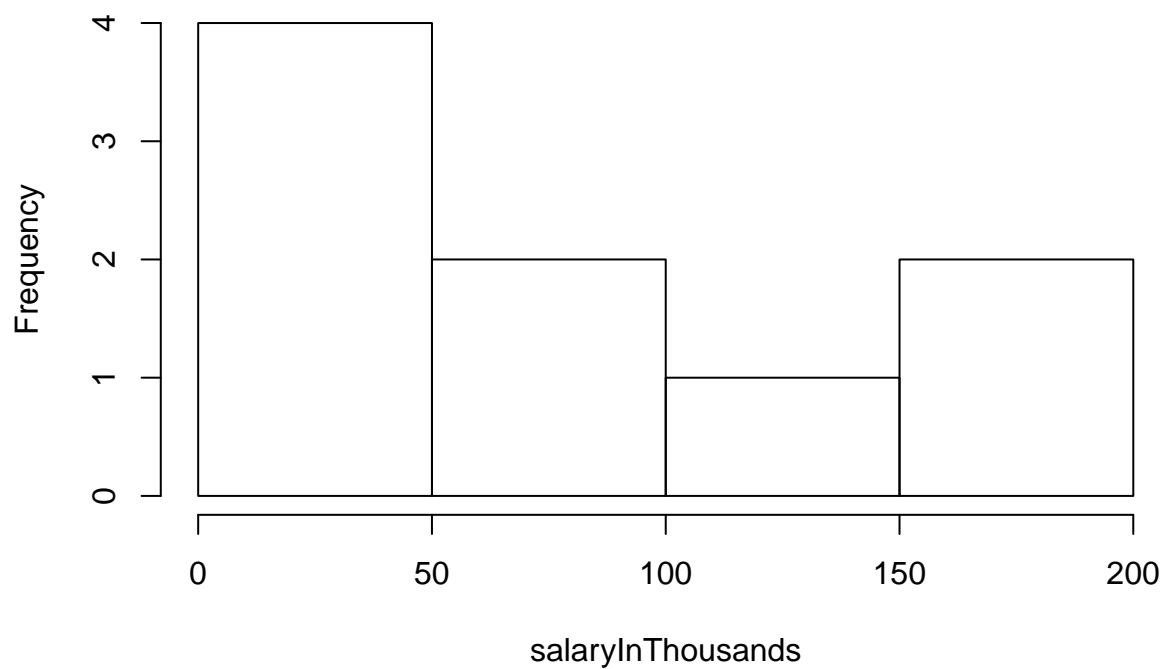
```
## [1] -0.05889984
```

```r
hist(age)
```
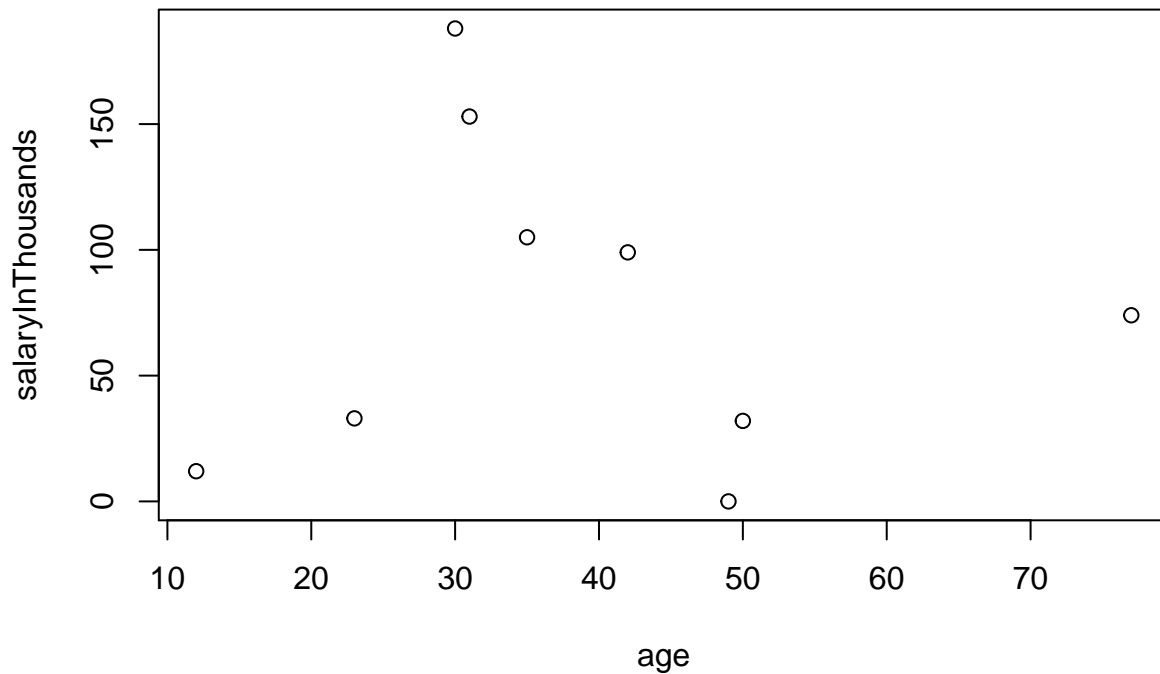
Histogram of age


```r
hist(salaryInThousands)
```

Histogram of salaryInThousands

```r
plot(age,salaryInThousands)
```



```r
# detach("package:PKGNAME", unload=TRUE)
# use library(PKGNAME) instead of require(PKGNAME)


doubleAtomicVector <- c(1, 3.14, 99.999)
integerAtomicVector <- c(1L, 3L, 19L)
logicalAtomicVector <- c(TRUE, FALSE, T, F)
characterAtomicVector <- c("this", "is a", "string")


myFavNum <- 3/10 # double
myNums <- c(1, 2, 3, 5, 8, 13, 21) # double
firstNames <- c("Dereck", "Nick") # character
myVec <- c("Perez", 25) # character

(myAtomicVector <- c(1, 2, 3, 4, -99, 5, NA, 4, 22.223))
```

```
## [1]   1.000   2.000   3.000   4.000 -99.000   5.000      NA   4.000  22.223
```

```r
myAtomicVector[5]
```

```
## [1] -99
```

```r
myAtomicVector[c(1,2,5,9)]
```

```
## [1]   1.000   2.000 -99.000  22.223
```

```r
myAtomicVector[10]
```

```
## [1] NA
```

```r
myAtomicVector[3:8]
```

```
## [1]   3   4 -99   5  NA   4
```

```r
myFirstAddition <- myNums[7] + myFavNum
mySecondAddition <- myNums + myFavNum
myFirstSum <- sum(c(myNums, myFavNum))
thisIsGettingMoreComplex <- min(myNums) + myFavNum
whatTypeOfVectorIsThis <- as.numeric(myVec[2]) + myNums[2]


(myAtomicVector_01 <- c(99.1, 98.2, 97.3, 96.4, NA))
```

```
## [1] 99.1 98.2 97.3 96.4   NA
```

```r
sum(myAtomicVector_01)
```

```
## [1] NA
```

```r
mean(myAtomicVector_01)
```

```
## [1] NA
```

```r
sum(myAtomicVector_01, na.rm = TRUE)
```

```
## [1] 391
```

```r
mean(myAtomicVector_01, na.rm = TRUE)
```

```
## [1] 97.75
```

```r
myVec <- c(1, NA)
myVec[2] == NA # NEVER DO THIS!
```

```
## [1] NA
```

```r
is.na(myVec[2])
```

```
## [1] TRUE
```

```
data() # AirPassengers
```

```
(myAtomicVector_01 <- c(99.1, 98.2, 97.3, 96.4))
```

```
## [1] 99.1 98.2 97.3 96.4
```

```
myAtomicVector_01[myAtomicVector_01 > 98] # equivalent to myAtomicVector_01[c(T, T, F, F)]
```

```
## [1] 99.1 98.2
```

```
myAtomicVector <- c(1, 4, 3, 2, NA, 3.22, -44, 2, NA, 0, 22, 34)
```

```
length(myAtomicVector[myAtomicVector>0 & !is.na(myAtomicVector)])
```

```
## [1] 8
```

```
length(myAtomicVector[myAtomicVector<0 & !is.na(myAtomicVector)])
```

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

```
length(myAtomicVector[myAtomicVector==0 & !is.na(myAtomicVector)])
```

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

```
length(myAtomicVector[is.na(myAtomicVector)])
```

```
## [1] 2
```

```
length(myAtomicVector[myAtomicVector!=0 & !is.na(myAtomicVector)])
```

```
## [1] 9
```

```
sum(myAtomicVector[myAtomicVector>0 & !is.na(myAtomicVector)])
```

```
## [1] 71.22
```

```
sum(myAtomicVector[myAtomicVector<0 & !is.na(myAtomicVector)])
```

```
## [1] -44
```

```
# Solution
sum(myAtomicVector > 0, na.rm = T)
```

```
## [1] 8
```

```r
sum(myAtomicVector < 0, na.rm = T)
```

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

```r
sum(myAtomicVector == 0, na.rm = T)
```

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

```r
sum(is.na(myAtomicVector))
```

```
## [1] 2
```

```r
sum(myAtomicVector != 0 & !is.na(myAtomicVector), na.rm = T) # sum(myAtomicVector != 0, na.rm = T)
```

```
## [1] 9
```

```r
sum(myAtomicVector[myAtomicVector > 0], na.rm = T)
```

```
## [1] 71.22
```

```r
sum(myAtomicVector[myAtomicVector < 0], na.rm = T)
```

```
## [1] -44
```

```r
(x <- c(1, NULL, 3))
```

```
## [1] 1 3
```

```r
x <- factor(c("M","F","F","M"))
bodyType <- factor(c("healthy", "healthy", "healthy", "obese", "overweight",
                     "overweight", "skinny"))
bodyType
```

```
## [1] healthy    healthy    healthy    obese      overweight overweight
## [7] skinny
## Levels: healthy obese overweight skinny
```

```r
levels(bodyType)
```

```
## [1] "healthy"    "obese"      "overweight" "skinny"
```

```r
bodyType < "obese"
```

```
## Warning in Ops.factor(bodyType, "obese"): '<' not meaningful for factors
```

```
## [1] NA NA NA NA NA NA NA
```

```r
(bodyType <- factor(c("h", "h", "h", "ob", "ov", "ov", "s"),
                     levels = c("s", "h", "ov", "ob"),
                     labels = c("Skinny", "Healthy", "Overweight", "Obese"),
                     ordered = TRUE))
```

```
## [1] Healthy    Healthy    Healthy    Obese      Overweight Overweight
## [7] Skinny
## Levels: Skinny < Healthy < Overweight < Obese
```

```r
levels(bodyType)
```

```
## [1] "Skinny"     "Healthy"     "Overweight" "Obese"
```

```r
bodyType < "Obese"
```

```
## [1]  TRUE  TRUE  TRUE FALSE  TRUE  TRUE  TRUE
```

```r
bodyType[bodyType < "Obese"]
```

```
## [1] Healthy    Healthy    Healthy    Overweight Overweight Skinny
## Levels: Skinny < Healthy < Overweight < Obese
```

```r
myCyl <- mtcars$cyl
ord <- factor(myCyl, levels = sort(unique(myCyl)), labels = c("Small", "Medium", "Large"), ordered = T)
sum(ord<="Medium")
```

```
## [1] 18
```

```r
(xyz <- data.frame(numberColumn = 1:3, letterColumn = c("a","b","c"),
                   stringsAsFactors = F))
```

```
##   numberColumn letterColumn
## 1            1            a
## 2            2            b
## 3            3            c
```

```r
str(xyz)
```

```
## 'data.frame':    3 obs. of  2 variables:
##  $ numberColumn: int  1 2 3
##  $ letterColumn: chr  "a" "b" "c"
```