R Basics

Ed Gonzalez

2023-01-03

```
##
## The downloaded binary packages are in
##  /var/folders/tz/sh20cj15711657_9_1d4v6m00000gn/T//Rtmp2VuaOO/downloaded_packages
```

```
summary(TitanicSurvival)
```

```
##  survived      sex           age          passengerClass
##  no :809   female:466   Min.   : 0.1667   1st:323
##  yes:500   male  :843   1st Qu.:21.0000   2nd:277
##                         Median :28.0000   3rd:709
##                         Mean   :29.8811
##                         3rd Qu.:39.0000
##                         Max.   :80.0000
##                         NA's   :263
```

Using '?' to find information about a function and a quick method of creating a table with specific variables

```
?table()
```

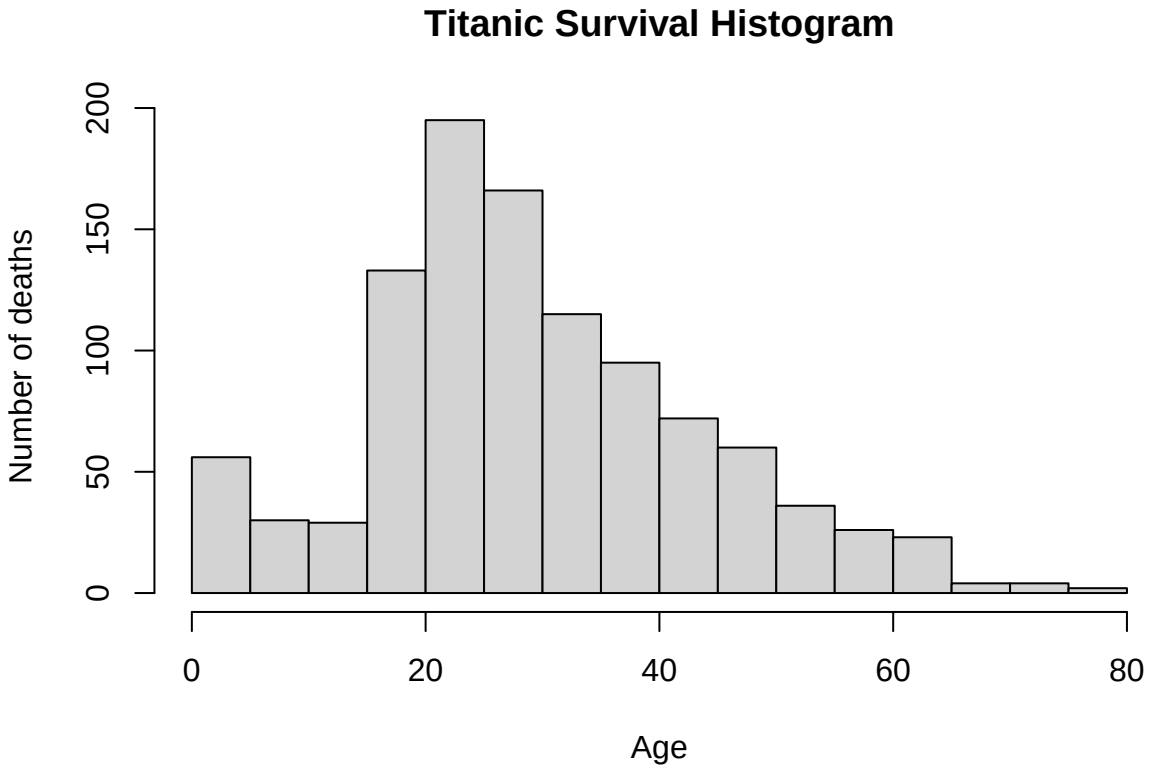
```
## Help on topic 'table' was found in the following packages:
##
##   Package               Library
##   vctrs                 /Library/Frameworks/R.framework/Versions/4.2-arm64/Resources/library
##   base                  /Library/Frameworks/R.framework/Resources/library
##
##
## Using the first match ...
```

```
table(TitanicSurvival$passengerClass, TitanicSurvival$survived)
```

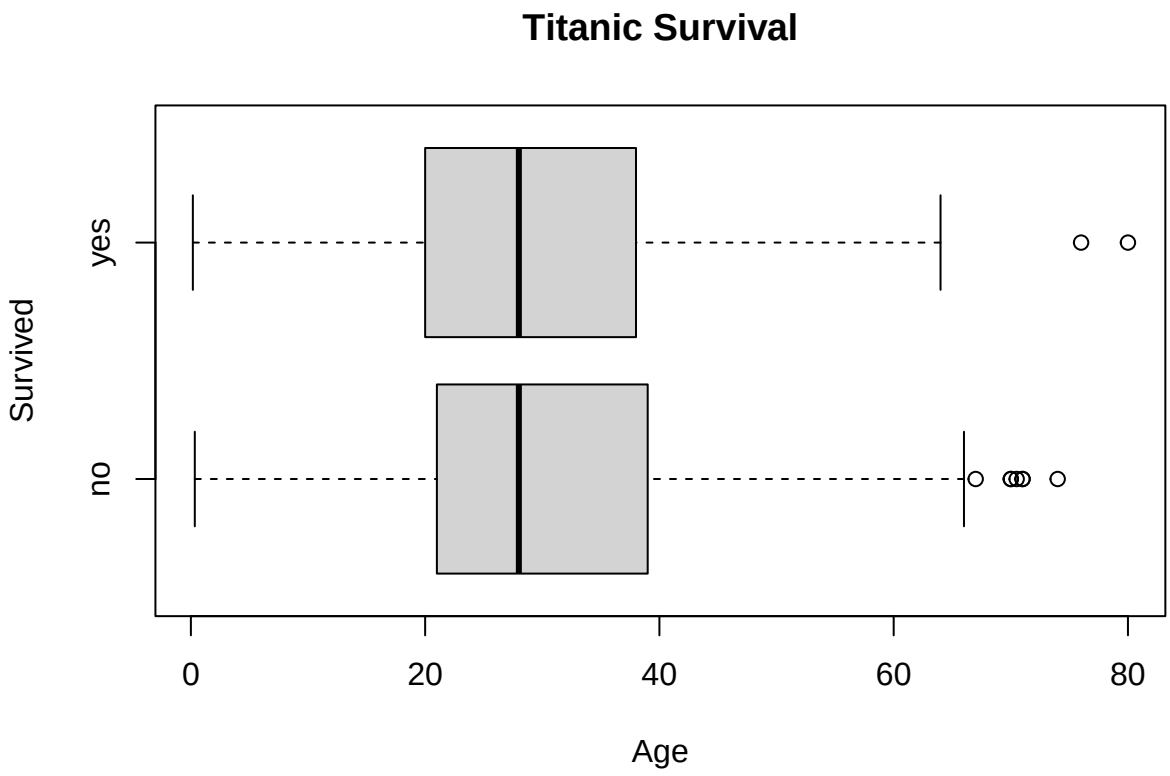
```
##
##        no yes
##   1st 123 200
##   2nd 158 119
##   3rd 528 181
```

Creating a histogram and boxplot using Base R functions

```
hist(TitanicSurvival$age,
     main = "Titanic Survival Histogram",
     xlab = "Age",
     ylab = "Number of deaths")
```



```
boxplot(age ~ survived, data = TitanicSurvival,
        main = "Titanic Survival",
        xlab = "Age",
        ylab = "Survived", horizontal = TRUE)
```



Downloading a dataset and assigning variable names

```
abalone.url <- "https://archive.ics.uci.edu/ml/machine-learning-databases/abalone/abalone.data"
abalone.data <- read.csv(abalone.url, header=FALSE)
names(abalone.data) <- c("sex", "length", "diameter",
                         "height", "weight.whole", "weight.shucked",
                         "weight.viscera", "weight.shell", "rings")
```

Functions that can identify the information contained within the data set, as well as creating variables

```
class(abalone.data)
```

```
## [1] "data.frame"
```

```
typeof("diameter")
```

```
## [1] "character"
```

```
summary(abalone.data$diameter[abalone.data$rings])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.2550  0.3700  0.3800  0.3838  0.4250  0.4800
```

```
mean(abalone.data$diameter[abalone.data$sex == "F"])
```

```
## [1] 0.4547322
```

```
x <- abalone.data$rings
x_t <- t(abalone.data$rings)
y <-abalone.data$diameter
```

Doing some arithmetic with the variables we created

```
solve(x_t%*%x)*(x_t%*%y)
```

```
##            [,1]
## [1,] 0.03883339
```