r_int_day_1_r_basic_sumstat

Nicholus Tint Zaw

2022-10-27

Content

r - base commands for Summary Statistics

```r
# load r-bulit-in dataset

df <- iris

# datafile summary
summary(df)
```

```
##   Sepal.Length    Sepal.Width     Petal.Length    Petal.Width
## Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100
## 1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300
## Median :5.800   Median :3.000   Median :4.350   Median :1.300
## Mean   :5.843   Mean   :3.057   Mean   :3.758   Mean   :1.199
## 3rd Qu.:6.400   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.800
## Max.   :7.900   Max.   :4.400   Max.   :6.900   Max.   :2.500
##        Species
## setosa    :50
## versicolor:50
## virginica :50
##
##
##
```

```r
# mean
mean(df$Sepal.Length)
```

```
## [1] 5.843333
```

```r
# SD
sd(df$Sepal.Length)
```

```
## [1] 0.8280661
```

```r
# median
median(df$Sepal.Length)
```

```
## [1] 5.8
```

```r
# percentile
quantile(df$Sepal.Width, 0.1)
```

```
## 10%
## 2.5
```

```r
quantile(df$Petal.Length, c(0.1, 0.9))
```

```
## 10% 90%
## 1.4 5.8
```

```r
quantile(df$Petal.Length, seq(0, 1, by = 0.1))
```

```
##   0%  10%  20%  30%  40%  50%  60%  70%  80%  90% 100%
## 1.00 1.40 1.50 1.70 3.90 4.35 4.64 5.00 5.32 5.80 6.90
```

```r
# quartile
quantile(df$Petal.Length, seq(0, 1, by = 0.25))
```

```
##   0%  25%  50%  75% 100%
## 1.00 1.60 4.35 5.10 6.90
```

```r
# IQR
IQR(df$Petal.Width)
```

```
## [1] 1.5
```

```r
quantile(df$Petal.Width, 0.75) - quantile(df$Petal.Width, 0.25)
```

```
## 75%
## 1.5
```

loading data files

```r
# load csv
df <- read.csv("wnba.csv")


# excel file
library("readxl")

df_2 <- read_excel("wnba.xlsx", sheet = "wnba") # want to learn more!
```

exercises

In a Science test, each correct answer was awarded 5 marks and 2 marks were deducted for each incorrect answer. If there were 20 questions in the test and Danny obtained 79 marks, how many questions did he answer correctly?

Let assume;

- T as total questions(or answers)
- C as total correct answer
- IC as total incorrect answer
- DM as Danny's total mark
- PT as possible total marks (when all answers are correct)
- AWD as award scale
- DED as deduction scale

We can write above problem in to the following mathematics equations.

- $$T = C + IC$$

- $$C = T - IC$$

- $$IC = T - C$$

- $$PT = AWD \times T$$

- $$DM = (AWD \times C) + (DED \times IC)$$

- $$DM = (AWD \times C + (DED \times (T - C))$$

- $$DM = (AWD \times C) + (DEC \times C) - (DEC \times T)$$

- $$DM = C(AWD + DEC) - (DEC \times T)$$

- $$C = \frac{DM + (DEC \times T)}{(AWD + DEC)}$$

```
awd <- 5
ded <- -2
t <- 20
dm <- 79

c <- (dm + (-2 * t))/(awd + ded)
c
```

```
## [1] 13
```

Compare with below method and any comment?

8 In a Science test, each correct answer was awarded 5 marks and 2 marks
deducted for each incorrect answer. If there were 20 questions in the test
Danny obtained 79 marks, how many questions did he answer correctly?

All question weve answered covrectly

20×5 = 100
100 - 79 = 21
Each wrong answer (5+2) = 7
         21 ÷ 7 = 3
         20 - 3 = 17