Stat 243 – Homework 1

Sharon Velpula

January 12, 2022

Importing dataset

```
file <- read.csv("https://scofield.site/teaching/data/csv/ssurv.csv")
head(file)
```

```
##   sex class gpa height pulse childrank numchildren haircut randomnum
## 1   F    So 3.6     NA    NA         2           3    2.00         6
## 2   F    So 3.4     NA    NA         4           4   10.00         7
## 3   M    Fr 3.0     71    68         2           4    0.00        17
## 4   M    So 2.6     72   100         2           1   15.00         3
## 5   M    So 2.2     68   101         4           3   11.00        13
## 6   M    So 2.4     72    74         2           2    9.99        11
##   speedtickets cds smoker hourssleep selfhandedness momhandedness dadhandedness
## 1            0  15    Non        8.0              R             L             R
## 2            0  10    Non        8.0              R             R             R
## 3            2  53    Non        5.5              R             R             R
## 4            3 170  Smoke        7.0              L             R             R
## 5            0  55  Smoke        8.0              R             R             R
## 6            0 101    Non        6.0              R             R             R
##     region oncampus cupscoffee birthday overtwenty
## 1 Suburban        Y          1       Th          Y
## 2    Rural        Y          0       Fr          N
## 3 Suburban        Y          0       Th          N
## 4 Suburban        Y          2       We          Y
## 5 Suburban        Y          0       Mo          N
## 6 Suburban        Y          0       Th          N
```

Identification of variables

Categorical variable: sex

Discrete quantitative variable: numchildren

Continuous quantitative variable: gpa

Command

```
filter(subset(file, select = c(sex, numchildren, gpa, smoker)), smoker == "Smoke")
```

```
##    sex numchildren   gpa smoker
## 1    M           1 2.600  Smoke
```

```
## 2    M           3 2.200  Smoke
## 3    M           4 2.987  Smoke
## 4    M           7    NA  Smoke
## 5    M           4 2.900  Smoke
## 6    M           4 3.300  Smoke
## 7    F           4 3.470  Smoke
## 8    M           3 3.750  Smoke
## 9    M           5 3.300  Smoke
## 10   M           4 3.400  Smoke
## 11   M           3 3.500  Smoke
## 12   F           2 2.500  Smoke
## 13   F           4 3.260  Smoke
## 14   M           5    NA  Smoke
## 15   M           4 1.300  Smoke
## 16   M           4 2.700  Smoke
## 17   F           2 2.336  Smoke
## 18   F           2 2.600  Smoke
```

Standard Deviation Formula

s=

$$\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2}$$