Exercise 5

Mingcong Pan

11/5/2021

Functions

- Create a function for the mean, median, and standard deviation.

```
mean <- function(x) {
 sum(x) / (length(x) - 1)
}

median <- function(x) {
  n <- length(x)
  s <- sort(x)
  ifelse(n%%2==1,s[(n+1)/2],mean(s[n/2+0:1]))
}

sd <- function(x) {
  sum((x - mean(x))^2) / (length(x) - 1)
}
```

- Create a function that finds the mean and excludes the lowest and highest value.

```
mean2 <- function(x) {
  mean(x[x > min(x) & x < max(x)])
}
```

- Apply the functions to the Christianity variables.

```
df <- read.csv("national.csv")

lapply(df[4:9], mean)
lapply(df[4:9], median)
lapply(df[4:9], sd)
lapply(df[4:9], mean2)
```

- Write a function that lists all the unique years with more than 300,000 Christians in total.

```
years <- function(df) {
  unique(df$year[df$christianity_all > 300000])
}
```

Loops/apply

- Write a loop to find how many variables there are per observation.

```
sum = 0
for(i in 1:ncol(df)){
  sum = sum + 1
}
```

- Write a loop to find the mean number of Protestant Christians in each country (i.e., the state column). Then use an apply family function to do the same.

```
tapply(df$christianity_all, df$state, mean)
```

- Check the column type for each variable.

```
for(i in 1:ncol(df)){
  typeof(df[,i])
}
```