

Exercise 5: Functions and Loops

Dillon Laaker

10/28/2020

Functions

0. Create a data set to use for practice below

```
setwd("~/Library/Mobile Documents/com~apple~CloudDocs/1. PS materials/PS 811/Week5-loops and functions")  
  
library(tidyverse)  
  
## -- Attaching packages ----- tidyverse 1.3.0 --  
  
## v ggplot2 3.3.3      v purrr    0.3.4  
## v tibble   3.1.4      v dplyr    1.0.7  
## v tidyr    1.1.3      v stringr  1.4.0  
## v readr    1.3.1      vforcats  0.5.0  
  
## -- Conflicts ----- tidyverse_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag()   masks stats::lag()  
  
data<- read.csv("national.csv")
```

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

```
# when a is a variable (column in dataframe)  
  
myfunc1 <- function(a){  
  mean <- sum(a) / length(a)  
  return(mean)  
}  
  
myfunc2 <- function(a){  
  n <- length(a)  
  s <- sort (a)  
  ifelse (n%%2==1, s[(n+1)/2], mean(s[n/2], s[(n/2)+1]))  
}  
  
myfunc3 <- function(a){  
  v <- sum((a - (sum(a) / length(a)))^2) / (length(a) - 1)  
  sd <- sqrt(v)  
  return(sd)  
}
```

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

```
myfunc4 <- function(a){  
  s<- sort(a)
```

```

d <- s[-c(1, length(s))]
v<-mean(d)
return(v)
}

```

3. Apply the functions to the Christianity variables.

```

r1 <- myfunc1(data$christianity_all)
r2 <- myfunc2(data$christianity_all)
r3 <- myfunc3(data$christianity_all)
r4 <- myfunc4(data$christianity_all)

```

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

```

myfunc5 <- function(d, a, b){
  x1<- as.data.frame(d[a>300000, b])
  u<- unique(x1)
  return(u)
}

r5 <-myfunc5(data, data$christianity_all, 'year')

#manual check
r5_t<-data %>%
  filter(christianity_all > 300000)%>%
  select(year)
unique(r5_t$year)

## [1] 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995 2000 2005 2010

```

Loops/apply

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

```

m <- data.frame(matrix(ncol = 1, nrow = nrow(data)))

for (i in 1:nrow(data)){
  m[i,1] <- ncol(data)
}

```

2. 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.

```

l <- unique(data$state)
m <- data.frame(matrix(ncol = 2, nrow = nrow(data.frame(l)))) #ncol = number of column

for(i in 1: length(l)) {
  x<- data %>% filter(state == l[i]) %>%
    summarise(m_PC = mean(christianity_protestant)) %>%
    pull(m_PC)
  m[i, 1] <- l[i]
  m[i, 2] <- x
}

#manual check
check <- data.frame(data %>% group_by(state) %>%

```

```
summarise (m_PC = mean(christianity_protestant))

#Using apply function
m2<- data.frame(tapply(data$christianity_protestant, data$state, mean))
```

3. Check the column type for each variable.

```
coltype <- data.frame(matrix(ncol = 2, nrow = ncol(data)))

for (i in 1:ncol(data)) {
  c<- class(data[, i])
  coltype[i, 1] <- colnames(data)[i]
  coltype[i, 2] <- c
}
```