Code for QSS Chapter 1: Introduction

Kosuke Imai

First Printing

Section 1.1: Overview of the Book

Section 1.2: How to Use this Book

```r
install.packages("swirl") # install the package
library(swirl) # load the package
install_course_github("kosukeimai", "qss-swirl") # install the course
library(swirl)
swirl()
```

Section 1.3: Introduction to R

Section 1.3.1: Arithmetic Operations

```r
5 + 3
```

```
## [1] 8
```

```r
5 - 3
```

```
## [1] 2
```

```r
5 / 3
```

```
## [1] 1.666667
```

```r
5 ^ 3
```

```
## [1] 125
```

```r
5 * (10 - 3)
```

```
## [1] 35
```

```r
sqrt(4)
```

```
## [1] 2
```

Section 1.3.2: Objects

```r
result <- 5 + 3
result
```

```
## [1] 8
```

```r
print(result)
```

```
## [1] 8
```

```
result <- 5 - 3
result
```

```
## [1] 2
```

```
kosuke <- "instructor"
kosuke
```

```
## [1] "instructor"
```

```
kosuke <- "instructor and author"
kosuke
```

```
## [1] "instructor and author"
```

```
Result <- "5"
Result
```

```
## [1] "5"
```

```
result
```

```
## [1] 2
```

```
class(result)
```

```
## [1] "numeric"
```

```
Result
```

```
## [1] "5"
```

```
class(Result)
```

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

```
class(sqrt)
```

```
## [1] "function"
```

Section 1.3.3: Vectors

```
world.pop <- c(2525779, 3026003, 3691173, 4449049, 5320817, 6127700, 6916183)
world.pop
```

```
## [1] 2525779 3026003 3691173 4449049 5320817 6127700 6916183
```

```
pop.first <- c(2525779, 3026003, 3691173)
pop.second <- c(4449049, 5320817, 6127700, 6916183)
pop.all <- c(pop.first, pop.second)
pop.all
```

```
## [1] 2525779 3026003 3691173 4449049 5320817 6127700 6916183
```

```
world.pop[2]
```

```
## [1] 3026003
```

```
world.pop[c(2, 4)]
```

```
## [1] 3026003 4449049
```

```r
world.pop[c(4, 2)]
```

```
## [1] 4449049 3026003
```

```r
world.pop[-3]
```

```
## [1] 2525779 3026003 4449049 5320817 6127700 6916183
```

```r
pop.million <- world.pop / 1000
pop.million
```

```
## [1] 2525.779 3026.003 3691.173 4449.049 5320.817 6127.700 6916.183
```

```r
pop.rate <- world.pop / world.pop[1]
pop.rate
```

```
## [1] 1.000000 1.198047 1.461400 1.761456 2.106604 2.426063 2.738238
```

```r
pop.increase <- world.pop[-1] - world.pop[-7]
percent.increase <- (pop.increase / world.pop[-7]) * 100
percent.increase
```

```
## [1] 19.80474 21.98180 20.53212 19.59448 15.16464 12.86752
```

```r
percent.increase[c(1, 2)] <- c(20, 22)
percent.increase
```

```
## [1] 20.00000 22.00000 20.53212 19.59448 15.16464 12.86752
```

Section 1.3.4: Functions

```r
length(world.pop)
```

```
## [1] 7
```

```r
min(world.pop)
```

```
## [1] 2525779
```

```r
max(world.pop)
```

```
## [1] 6916183
```

```r
range(world.pop)
```

```
## [1] 2525779 6916183
```

```r
mean(world.pop)
```

```
## [1] 4579529
```

```r
sum(world.pop) / length(world.pop)
```

```
## [1] 4579529
```

```r
year <- seq(from = 1950, to = 2010, by = 10)
year
```

```
## [1] 1950 1960 1970 1980 1990 2000 2010
```

```r
seq(to = 2010, by = 10, from = 1950)
```

```
## [1] 1950 1960 1970 1980 1990 2000 2010
```

```r
seq(from = 2010, to = 1950, by = -10)
```

```
## [1] 2010 2000 1990 1980 1970 1960 1950
```

```r
2008:2012
```

```
## [1] 2008 2009 2010 2011 2012
```

```r
2012:2008
```

```
## [1] 2012 2011 2010 2009 2008
```

```r
names(world.pop)
```

```
## NULL
```

```r
names(world.pop) <- year
names(world.pop)
```

```
## [1] "1950" "1960" "1970" "1980" "1990" "2000" "2010"
```

```r
world.pop
```

```
##    1950    1960    1970    1980    1990    2000    2010 
## 2525779 3026003 3691173 4449049 5320817 6127700 6916183
```

```r
## myfunction <- function(input1, input2, ..., inputN) {
##
##     DEFINE `output' USING INPUTS
##
##     return(output)
## }

my.summary <- function(x){ # function takes one input
  s.out <- sum(x)
  l.out <- length(x)
  m.out <- s.out / l.out
  out <- c(s.out, l.out, m.out) # define the output
  names(out) <- c("sum", "length", "mean") # add labels
  return(out) # end function by calling output
}
z <- 1:10
my.summary(z)
```

```
##    sum length   mean 
##   55.0   10.0    5.5
```

```r
my.summary(world.pop)
```

```
##      sum   length     mean 
## 32056704        7  4579529
```

Section 1.3.5: Data Files

```
## setwd("qss/INTRO")
## getwd()
```

```
UNpop <- read.csv("UNpop.csv")
class(UNpop)
```

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

```
load("UNpop.RData")

names(UNpop)
```

```
## [1] "year"      "world.pop"
```

```
nrow(UNpop)
```

```
## [1] 7
```

```
ncol(UNpop)
```

```
## [1] 2
```

```
dim(UNpop)
```

```
## [1] 7 2
```

```
summary(UNpop)
```

```
##       year        world.pop
##  Min.   :1950   Min.   :2525779
##  1st Qu.:1965   1st Qu.:3358588
##  Median :1980   Median :4449049
##  Mean   :1980   Mean   :4579529
##  3rd Qu.:1995   3rd Qu.:5724258
##  Max.   :2010   Max.   :6916183
```

```
UNpop$world.pop
```

```
## [1] 2525779 3026003 3691173 4449049 5320817 6127700 6916183
```

```
UNpop[, "world.pop"] # extract the column called "world.pop"
```

```
## [1] 2525779 3026003 3691173 4449049 5320817 6127700 6916183
```

```
UNpop[c(1, 2, 3),]   # extract the first three rows (and all columns)
```

```
##   year world.pop
## 1 1950   2525779
## 2 1960   3026003
## 3 1970   3691173
```

```
UNpop[1:3, "year"]   # extract the first three rows of the "year" column
```

```
## [1] 1950 1960 1970
```

```
## take elements 1, 3, 5, ... of the "world.pop" variable
UNpop$world.pop[seq(from = 1, to = nrow(UNpop), by = 2)]
```

```
## [1] 2525779 3691173 5320817 6916183
```

```r
world.pop <- c(UNpop$world.pop, NA)
world.pop
```

```
## [1] 2525779 3026003 3691173 4449049 5320817 6127700 6916183      NA
```

```r
mean(world.pop)
```

```
## [1] NA
```

```r
mean(world.pop, na.rm = TRUE)
```

```
## [1] 4579529
```

Section 1.3.6: Saving Objects

```r
## save.image("qss/INTRO/Chapter1.RData")

## save(UNpop, file = "Chapter1.RData")
## save(world.pop, year, file = "qss/INTRO/Chapter1.RData")

## write.csv(UNpop, file = "UNpop.csv")

## load("Chapter1.RData")
```

Section 1.3.7: Packages

```r
install.packages("foreign") # install package
library("foreign") # load package

read.dta("UNpop.dta")
read.spss("UNpop.sav")

write.dta(UNpop, file = "UNpop.dta")
```

Section 1.3.8: Programming and Learning Tips

```r
source("UNpop.R")

##
## File: UNpop.R
## Author: Kosuke Imai
## The code loads the UN population data and saves it as a STATA file
##

library(foreign)
UNpop <- read.csv("UNpop.csv")
UNpop$world.pop <- UNpop$world.pop / 1000  # population in millions
write.dta(UNpop, file = "UNpop.dta")

library(lintr)
lint("UNpop.R")
```