

Pre-Class Work for September 27

Maggie Matsui

Question 1:

Using a loop, print the integers from 1 to 50.

```
for(i in 1:50){  
  print(i)  
}
```

```
## [1] 1  
## [1] 2  
## [1] 3  
## [1] 4  
## [1] 5  
## [1] 6  
## [1] 7  
## [1] 8  
## [1] 9  
## [1] 10  
## [1] 11  
## [1] 12  
## [1] 13  
## [1] 14  
## [1] 15  
## [1] 16  
## [1] 17  
## [1] 18  
## [1] 19  
## [1] 20  
## [1] 21  
## [1] 22  
## [1] 23  
## [1] 24  
## [1] 25  
## [1] 26  
## [1] 27  
## [1] 28  
## [1] 29  
## [1] 30  
## [1] 31  
## [1] 32  
## [1] 33  
## [1] 34  
## [1] 35  
## [1] 36  
## [1] 37  
## [1] 38  
## [1] 39  
## [1] 40  
## [1] 41
```

```
## [1] 42
## [1] 43
## [1] 44
## [1] 45
## [1] 46
## [1] 47
## [1] 48
## [1] 49
## [1] 50
```

Question 2:

A. Using a loop, add all the integers between 0 and 1000.

```
sum <- 0
for(i in 0:1000){
  sum <- sum + i
}
sum
```

```
## [1] 500500
```

B. Now, add all the EVEN integers between 0 and 1000 (hint: use seq())

```
sum <- 0
for(i in seq(0, 1000, by = 2)){
  sum <- sum + i
}
sum
```

```
## [1] 250500
```

C. Now, repeat A and B WITHOUT using a loop.

```
sum(c(0:1000))
```

```
## [1] 500500
```

```
sum(seq(0, 1000, by = 2))
```

```
## [1] 250500
```

Question 3:

Here is a dataframe of survey data containing 5 questions :

```
survey <- data.frame(
  "participant" = c(1, 2, 3, 4, 5, 6),
  "q1" = c(5, 3, 2, 7, 11, 0),
  "q2" = c(4, 2, 2, 5, -10, 99),
  "q3" = c(-4, -3, 4, 2, 9, 10),
  "q4" = c(-30, 5, 2, 23, 4, 2),
  "q5" = c(88, 4, -20, 2, 4, 2)
)
```

The response to each question should be an integer between 1 and 5. Obviously, we have some bad values in the dataframe. The goal of this problem is to fix them.

A. Using a loop, create a new dataframe called `survey.clean` where all the invalid values (those that are not integers between 1 and 5) are set to NA.

```
survey.clean <- survey
for(i in 1:dim(survey.clean)[1]){
  for(j in 1:dim(survey.clean)[2]){
    if(survey.clean[i,j] < 1 | survey.clean[i,j] > 5){
      survey.clean[i,j] <- NA
    }
  }
}

survey.clean
```

```
## participant q1 q2 q3 q4 q5
## 1          1  5  4 NA NA NA
## 2          2  3  2 NA  5  4
## 3          3  2  2  4  2 NA
## 4          4 NA  5  2 NA  2
## 5          5 NA NA NA  4  4
## 6         NA NA NA NA  2  2
```

B. Now, again using a loop, add a new column to the dataframe called “invalid.answers” that indicates, for each participant, how many bad answers they gave.

```
survey.clean[,dim(survey.clean)[2] + 1] <- numeric(dim(survey.clean)[1])
for(i in 1:dim(survey.clean)[1]){
  survey.clean[i, dim(survey.clean)[2]] <- sum(is.na(survey.clean[i,]))
}
names(survey.clean)[dim(survey.clean)[2]] <- "invalid.answers"

survey.clean
```

```
## participant q1 q2 q3 q4 q5 invalid.answers
## 1          1  5  4 NA NA NA              3
## 2          2  3  2 NA  5  4              1
## 3          3  2  2  4  2 NA              1
## 4          4 NA  5  2 NA  2              2
## 5          5 NA NA NA  4  4              3
## 6         NA NA NA NA  2  2              4
```