null

Question 1:

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

```
x<-1:50
for(i in length(x)){
  print(as.integer(x))
}
```

```
##  [1]  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
## [24] 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46
## [47] 47 48 49 50
```

Question 2:

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

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

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

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

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

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

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

```
sum(1:1000)
```

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

```
sum(seq(1,1000,2))
```

```
## [1] 250000
```

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 <- 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)
                    )

survey.clean<- survey
for(i in 2:ncol(survey.clean)) {
  data.col <- survey.clean[,i]
 data.col[(data.col %in% 1:5) == F]<-NA
 survey.clean[,i] <- data.col
}
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           6 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$invalid.answers<-NA
for(row.i in 1: nrow(survey.clean)) {
  data.col <- survey.clean[row.i,]
  n.na <- sum(is.na(data.col)) - 1
  survey.clean$invalid.answers[row.i] <- n.na
  }
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           6 NA NA NA  2  2               3
```