Solutions

Second assignment

Random vector

1. Generate a random normal vector of size 100
2. Compute its mean with for/repeat loop
3. Compute its variance withfor/repeat loop

```
vect = rnorm(100)
sum_vect = 0
for (i in vect){
sum_vect = sum_vect + i
}
mean_vect = sum_vect/length(vect)
mean_vect
sum_vect2 = 0
for (j in vect){
sum_vect2= sum_vect2+ (j-mean_vect)^2 }
var_vect=sum_vect2/(length(vect)-1)
var_vect
```

Second assignment

missing values

1. Use the airquality dataset from base
2. Compute the percentage p_na of missing values in a column
3. If p_na > 0,5 □ delete the column
4. If p_na <= 0,5 □ replace the missing values by 0 or by the mean of the column, depending on a variable "type_na"

```
for (col in colnames(airquality)) {
  s_na <- sum(is.na(airquality[[col]]))
  len <- length(airquality[[col]])
  p_na <- s_na / len
  p_na
```

```
airquality1 = airquality[,!(p_na > 0.05)]
```

```
replace_na = function(data, na) {
if (na == "0"){
for (i in 1:length(data)){
data[i][is.na(data[i])] = 0 }
} else if (na == "mean") {
for (i in 1:length(data)){
data[i][is.na(data[i])] = mean(data[,i],na.rm = TRUE)
}
}
return(data)
}
replace_na(data=airquality1, na = "0")
```

Second assignment

mean and standard deviation over the columns

1. Compute the mean of all columns of iris dataset
2. Compute their standard deviation

```
data(iris)
Mean_iris <- apply(select(iris, -(Species)), 2, mean)
Std_iris <- apply(select(iris, -(Species)), 2, sd)
```

```
colMeans(iris[1:4])
sapply(iris[1:4], function(x) sd(x))
```