Air Pollution

library(rio)

## Warning: package 'rio' was built under R version 3.6.2

library(moments)  
  
# Function, 'pollutantmean', calculating mean of a pollutant (sulfate or nitrate).   
m=0  
l=0  
files = list.files(pattern = "\*.csv")  
pollutantmean = function(directory, pollutant, id=1:332){  
 for(i in id[1]:id[length(id)])  
 {  
 data=0  
 data=import(files[i])  
 if(pollutant=="nitrate"){  
 m=m+sum(data$nitrate, na.rm=TRUE)  
 l = l + (length(data$nitrate) - sum(is.na(data$nitrate)))  
 }  
 else  
 {  
 m=m+sum(data$sulfate, na.rm=TRUE)  
 l = l + (length(data$nitrate) - sum(is.na(data$nitrate)))  
 }  
 }  
 m/l  
}  
  
# Number of completely observed cases in each data file.  
j=1  
df=data.frame()  
complete<-function(directory,id=1:332){  
 for(i in id[1]:id[length(id)])  
 {  
 data=0  
 data=import(files[i])  
 df[j,1]=i  
 df[j,2]=sum(complete.cases(data))   
 j=j+1   
 }  
 df  
}  
  
# Correlation between pollutant where number of completely cases is greater than threshold.  
j=1  
vec=c()  
corr<- function(directory, threshold=0){  
 for(i in 1:332){  
 data=0  
 data=import(files[i])  
 if(sum(complete.cases(data))>threshold){  
 vec[j]=cor(data$sulfate,data$nitrate,use = "complete.obs")   
 j=j+1  
 }  
 }  
 vec  
}