```
  library(VIM) # function aggr: visualize the missing value
library(tidyverse) #To use ggplot2, tidyr, dplyr
library(plotly) #To create interactive plots
library(DT) #To display the data
library(magrittr) #To pipe operators
library(ggplot2) #To make and customize quickly plots
library(devtools) #To Make Developing R Packages Easier
library(lubridate) # date tranformation
library(beginr)

beijing.data <- read.csv("PRSA_data_2010.1.1-2014.12.31.csv", header = T) # load the data set
head(beijing.data)
tail(beijing.data)

sum(is.na(beijing.data))
aggr(beijing.data, prop = T, number = T)

i <- NULL
j <- 1
compare_value_j <- 1
for ( i in 2010:2014){
  data_i <- beijing.data[beijing.data$year == i,]
  if (compare_value_j < length(na.omit(data_i$pm2.5))){
    compare_value_j <- length(na.omit(data_i$pm2.5))
    j <- j + 1
  }
  print(j + 2009) # the year will least missing value
}
beijing.data <- as_tibble(beijing.data)

i <- NULL
for ( i in 1:length(beijing.data$No)){
  if(beijing.data$month[i] == 3){
    beijing.data$season[i] <- 1
  }
  if(beijing.data$month[i] == 4){
    beijing.data$season[i] <- 1
  }
  if(beijing.data$month[i] == 5){
    beijing.data$season[i] <- 1
  }
  if(beijing.data$month[i] == 6){
    beijing.data$season[i] <- 2
  }
  if(beijing.data$month[i] == 7){
    beijing.data$season[i] <- 2
  }
  if(beijing.data$month[i] == 8){
    beijing.data$season[i] <- 2
  }
  if(beijing.data$month[i] == 9){
    beijing.data$season[i] <- 3
  }
  if(beijing.data$month[i] == 10){
    beijing.data$season[i] <- 3
  }
  if(beijing.data$month[i] == 11){
    beijing.data$season[i] <- 3
  }
  if(beijing.data$month[i] == 12){
    beijing.data$season[i] <- 4
  }
  if(beijing.data$month[i] == 1){
    beijing.data$season[i] <- 4
  }
  if(beijing.data$month[i] == 2){
    beijing.data$season[i] <- 4
  }
}
head(beijing.data)

cleanbeijing <-select(beijing.data, c("year","month","day","hour","season","pm2.5","cbwd","Iws", "Is
  na.omit() %>%
  filter(year >= 2013)%>%
  unite(timebyday, c("year", "month", "day"), remove = FALSE, sep = "-")
datatable(cleanbeijing, option = list(scrollX = TRUE))
```