Untitled

lee6

7/18/2021

ju<-airquality[complete.cases(airquality),]  
str(ju)

## 'data.frame': 111 obs. of 6 variables:  
## $ Ozone : int 41 36 12 18 23 19 8 16 11 14 ...  
## $ Solar.R: int 190 118 149 313 299 99 19 256 290 274 ...  
## $ Wind : num 7.4 8 12.6 11.5 8.6 13.8 20.1 9.7 9.2 10.9 ...  
## $ Temp : int 67 72 74 62 65 59 61 69 66 68 ...  
## $ Month : int 5 5 5 5 5 5 5 5 5 5 ...  
## $ Day : int 1 2 3 4 7 8 9 12 13 14 ...

#in the above output you notice that the data.frame contains only 111 obs #how to count the number of missing values in the airquality dataset

sum(is.na(airquality))

## [1] 44

summary(ju)

## Ozone Solar.R Wind Temp   
## Min. : 1.0 Min. : 7.0 Min. : 2.30 Min. :57.00   
## 1st Qu.: 18.0 1st Qu.:113.5 1st Qu.: 7.40 1st Qu.:71.00   
## Median : 31.0 Median :207.0 Median : 9.70 Median :79.00   
## Mean : 42.1 Mean :184.8 Mean : 9.94 Mean :77.79   
## 3rd Qu.: 62.0 3rd Qu.:255.5 3rd Qu.:11.50 3rd Qu.:84.50   
## Max. :168.0 Max. :334.0 Max. :20.70 Max. :97.00   
## Month Day   
## Min. :5.000 Min. : 1.00   
## 1st Qu.:6.000 1st Qu.: 9.00   
## Median :7.000 Median :16.00   
## Mean :7.216 Mean :15.95   
## 3rd Qu.:9.000 3rd Qu.:22.50   
## Max. :9.000 Max. :31.00

g<-c(1,2,3,4,5,6,7,8,9)  
vbgh<-matrix(g,ncol = 3)  
vbgh

## [,1] [,2] [,3]  
## [1,] 1 4 7  
## [2,] 2 5 8  
## [3,] 3 6 9

#how to remove missing values using drop na()function

library(tidyr)

## Warning: package 'tidyr' was built under R version 3.6.3

viewdata<-airquality%>%drop\_na()  
viewdata

## Ozone Solar.R Wind Temp Month Day  
## 1 41 190 7.4 67 5 1  
## 2 36 118 8.0 72 5 2  
## 3 12 149 12.6 74 5 3  
## 4 18 313 11.5 62 5 4  
## 5 23 299 8.6 65 5 7  
## 6 19 99 13.8 59 5 8  
## 7 8 19 20.1 61 5 9  
## 8 16 256 9.7 69 5 12  
## 9 11 290 9.2 66 5 13  
## 10 14 274 10.9 68 5 14  
## 11 18 65 13.2 58 5 15  
## 12 14 334 11.5 64 5 16  
## 13 34 307 12.0 66 5 17  
## 14 6 78 18.4 57 5 18  
## 15 30 322 11.5 68 5 19  
## 16 11 44 9.7 62 5 20  
## 17 1 8 9.7 59 5 21  
## 18 11 320 16.6 73 5 22  
## 19 4 25 9.7 61 5 23  
## 20 32 92 12.0 61 5 24  
## 21 23 13 12.0 67 5 28  
## 22 45 252 14.9 81 5 29  
## 23 115 223 5.7 79 5 30  
## 24 37 279 7.4 76 5 31  
## 25 29 127 9.7 82 6 7  
## 26 71 291 13.8 90 6 9  
## 27 39 323 11.5 87 6 10  
## 28 23 148 8.0 82 6 13  
## 29 21 191 14.9 77 6 16  
## 30 37 284 20.7 72 6 17  
## 31 20 37 9.2 65 6 18  
## 32 12 120 11.5 73 6 19  
## 33 13 137 10.3 76 6 20  
## 34 135 269 4.1 84 7 1  
## 35 49 248 9.2 85 7 2  
## 36 32 236 9.2 81 7 3  
## 37 64 175 4.6 83 7 5  
## 38 40 314 10.9 83 7 6  
## 39 77 276 5.1 88 7 7  
## 40 97 267 6.3 92 7 8  
## 41 97 272 5.7 92 7 9  
## 42 85 175 7.4 89 7 10  
## 43 10 264 14.3 73 7 12  
## 44 27 175 14.9 81 7 13  
## 45 7 48 14.3 80 7 15  
## 46 48 260 6.9 81 7 16  
## 47 35 274 10.3 82 7 17  
## 48 61 285 6.3 84 7 18  
## 49 79 187 5.1 87 7 19  
## 50 63 220 11.5 85 7 20  
## 51 16 7 6.9 74 7 21  
## 52 80 294 8.6 86 7 24  
## 53 108 223 8.0 85 7 25  
## 54 20 81 8.6 82 7 26  
## 55 52 82 12.0 86 7 27  
## 56 82 213 7.4 88 7 28  
## 57 50 275 7.4 86 7 29  
## 58 64 253 7.4 83 7 30  
## 59 59 254 9.2 81 7 31  
## 60 39 83 6.9 81 8 1  
## 61 9 24 13.8 81 8 2  
## 62 16 77 7.4 82 8 3  
## 63 122 255 4.0 89 8 7  
## 64 89 229 10.3 90 8 8  
## 65 110 207 8.0 90 8 9  
## 66 44 192 11.5 86 8 12  
## 67 28 273 11.5 82 8 13  
## 68 65 157 9.7 80 8 14  
## 69 22 71 10.3 77 8 16  
## 70 59 51 6.3 79 8 17  
## 71 23 115 7.4 76 8 18  
## 72 31 244 10.9 78 8 19  
## 73 44 190 10.3 78 8 20  
## 74 21 259 15.5 77 8 21  
## 75 9 36 14.3 72 8 22  
## 76 45 212 9.7 79 8 24  
## 77 168 238 3.4 81 8 25  
## 78 73 215 8.0 86 8 26  
## 79 76 203 9.7 97 8 28  
## 80 118 225 2.3 94 8 29  
## 81 84 237 6.3 96 8 30  
## 82 85 188 6.3 94 8 31  
## 83 96 167 6.9 91 9 1  
## 84 78 197 5.1 92 9 2  
## 85 73 183 2.8 93 9 3  
## 86 91 189 4.6 93 9 4  
## 87 47 95 7.4 87 9 5  
## 88 32 92 15.5 84 9 6  
## 89 20 252 10.9 80 9 7  
## 90 23 220 10.3 78 9 8  
## 91 21 230 10.9 75 9 9  
## 92 24 259 9.7 73 9 10  
## 93 44 236 14.9 81 9 11  
## 94 21 259 15.5 76 9 12  
## 95 28 238 6.3 77 9 13  
## 96 9 24 10.9 71 9 14  
## 97 13 112 11.5 71 9 15  
## 98 46 237 6.9 78 9 16  
## 99 18 224 13.8 67 9 17  
## 100 13 27 10.3 76 9 18  
## 101 24 238 10.3 68 9 19  
## 102 16 201 8.0 82 9 20  
## 103 13 238 12.6 64 9 21  
## 104 23 14 9.2 71 9 22  
## 105 36 139 10.3 81 9 23  
## 106 7 49 10.3 69 9 24  
## 107 14 20 16.6 63 9 25  
## 108 30 193 6.9 70 9 26  
## 109 14 191 14.3 75 9 28  
## 110 18 131 8.0 76 9 29  
## 111 20 223 11.5 68 9 30

str(viewdata)

## 'data.frame': 111 obs. of 6 variables:  
## $ Ozone : int 41 36 12 18 23 19 8 16 11 14 ...  
## $ Solar.R: int 190 118 149 313 299 99 19 256 290 274 ...  
## $ Wind : num 7.4 8 12.6 11.5 8.6 13.8 20.1 9.7 9.2 10.9 ...  
## $ Temp : int 67 72 74 62 65 59 61 69 66 68 ...  
## $ Month : int 5 5 5 5 5 5 5 5 5 5 ...  
## $ Day : int 1 2 3 4 7 8 9 12 13 14 ...

data56<-airquality[complete.cases(airquality),]  
str(data56)

## 'data.frame': 111 obs. of 6 variables:  
## $ Ozone : int 41 36 12 18 23 19 8 16 11 14 ...  
## $ Solar.R: int 190 118 149 313 299 99 19 256 290 274 ...  
## $ Wind : num 7.4 8 12.6 11.5 8.6 13.8 20.1 9.7 9.2 10.9 ...  
## $ Temp : int 67 72 74 62 65 59 61 69 66 68 ...  
## $ Month : int 5 5 5 5 5 5 5 5 5 5 ...  
## $ Day : int 1 2 3 4 7 8 9 12 13 14 ...

#using filter function

data54<-na.omit(airquality)  
str(data54)

## 'data.frame': 111 obs. of 6 variables:  
## $ Ozone : int 41 36 12 18 23 19 8 16 11 14 ...  
## $ Solar.R: int 190 118 149 313 299 99 19 256 290 274 ...  
## $ Wind : num 7.4 8 12.6 11.5 8.6 13.8 20.1 9.7 9.2 10.9 ...  
## $ Temp : int 67 72 74 62 65 59 61 69 66 68 ...  
## $ Month : int 5 5 5 5 5 5 5 5 5 5 ...  
## $ Day : int 1 2 3 4 7 8 9 12 13 14 ...  
## - attr(\*, "na.action")= 'omit' Named int 5 6 10 11 25 26 27 32 33 34 ...  
## ..- attr(\*, "names")= chr "5" "6" "10" "11" ...

data566<-airquality[rowSums(is.na(airquality))==0, ]  
str(data566)

## 'data.frame': 111 obs. of 6 variables:  
## $ Ozone : int 41 36 12 18 23 19 8 16 11 14 ...  
## $ Solar.R: int 190 118 149 313 299 99 19 256 290 274 ...  
## $ Wind : num 7.4 8 12.6 11.5 8.6 13.8 20.1 9.7 9.2 10.9 ...  
## $ Temp : int 67 72 74 62 65 59 61 69 66 68 ...  
## $ Month : int 5 5 5 5 5 5 5 5 5 5 ...  
## $ Day : int 1 2 3 4 7 8 9 12 13 14 ...

str(iris)

## 'data.frame': 150 obs. of 5 variables:  
## $ Sepal.Length: num 5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...  
## $ Sepal.Width : num 3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...  
## $ Petal.Length: num 1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...  
## $ Petal.Width : num 0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...  
## $ Species : Factor w/ 3 levels "setosa","versicolor",..: 1 1 1 1 1 1 1 1 1 1 ...

library(zoo)

## Warning: package 'zoo' was built under R version 3.6.3

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

data4566<-na.aggregate(airquality)  
data4566

## Ozone Solar.R Wind Temp Month Day  
## 1 41.00000 190.0000 7.4 67 5 1  
## 2 36.00000 118.0000 8.0 72 5 2  
## 3 12.00000 149.0000 12.6 74 5 3  
## 4 18.00000 313.0000 11.5 62 5 4  
## 5 42.12931 185.9315 14.3 56 5 5  
## 6 28.00000 185.9315 14.9 66 5 6  
## 7 23.00000 299.0000 8.6 65 5 7  
## 8 19.00000 99.0000 13.8 59 5 8  
## 9 8.00000 19.0000 20.1 61 5 9  
## 10 42.12931 194.0000 8.6 69 5 10  
## 11 7.00000 185.9315 6.9 74 5 11  
## 12 16.00000 256.0000 9.7 69 5 12  
## 13 11.00000 290.0000 9.2 66 5 13  
## 14 14.00000 274.0000 10.9 68 5 14  
## 15 18.00000 65.0000 13.2 58 5 15  
## 16 14.00000 334.0000 11.5 64 5 16  
## 17 34.00000 307.0000 12.0 66 5 17  
## 18 6.00000 78.0000 18.4 57 5 18  
## 19 30.00000 322.0000 11.5 68 5 19  
## 20 11.00000 44.0000 9.7 62 5 20  
## 21 1.00000 8.0000 9.7 59 5 21  
## 22 11.00000 320.0000 16.6 73 5 22  
## 23 4.00000 25.0000 9.7 61 5 23  
## 24 32.00000 92.0000 12.0 61 5 24  
## 25 42.12931 66.0000 16.6 57 5 25  
## 26 42.12931 266.0000 14.9 58 5 26  
## 27 42.12931 185.9315 8.0 57 5 27  
## 28 23.00000 13.0000 12.0 67 5 28  
## 29 45.00000 252.0000 14.9 81 5 29  
## 30 115.00000 223.0000 5.7 79 5 30  
## 31 37.00000 279.0000 7.4 76 5 31  
## 32 42.12931 286.0000 8.6 78 6 1  
## 33 42.12931 287.0000 9.7 74 6 2  
## 34 42.12931 242.0000 16.1 67 6 3  
## 35 42.12931 186.0000 9.2 84 6 4  
## 36 42.12931 220.0000 8.6 85 6 5  
## 37 42.12931 264.0000 14.3 79 6 6  
## 38 29.00000 127.0000 9.7 82 6 7  
## 39 42.12931 273.0000 6.9 87 6 8  
## 40 71.00000 291.0000 13.8 90 6 9  
## 41 39.00000 323.0000 11.5 87 6 10  
## 42 42.12931 259.0000 10.9 93 6 11  
## 43 42.12931 250.0000 9.2 92 6 12  
## 44 23.00000 148.0000 8.0 82 6 13  
## 45 42.12931 332.0000 13.8 80 6 14  
## 46 42.12931 322.0000 11.5 79 6 15  
## 47 21.00000 191.0000 14.9 77 6 16  
## 48 37.00000 284.0000 20.7 72 6 17  
## 49 20.00000 37.0000 9.2 65 6 18  
## 50 12.00000 120.0000 11.5 73 6 19  
## 51 13.00000 137.0000 10.3 76 6 20  
## 52 42.12931 150.0000 6.3 77 6 21  
## 53 42.12931 59.0000 1.7 76 6 22  
## 54 42.12931 91.0000 4.6 76 6 23  
## 55 42.12931 250.0000 6.3 76 6 24  
## 56 42.12931 135.0000 8.0 75 6 25  
## 57 42.12931 127.0000 8.0 78 6 26  
## 58 42.12931 47.0000 10.3 73 6 27  
## 59 42.12931 98.0000 11.5 80 6 28  
## 60 42.12931 31.0000 14.9 77 6 29  
## 61 42.12931 138.0000 8.0 83 6 30  
## 62 135.00000 269.0000 4.1 84 7 1  
## 63 49.00000 248.0000 9.2 85 7 2  
## 64 32.00000 236.0000 9.2 81 7 3  
## 65 42.12931 101.0000 10.9 84 7 4  
## 66 64.00000 175.0000 4.6 83 7 5  
## 67 40.00000 314.0000 10.9 83 7 6  
## 68 77.00000 276.0000 5.1 88 7 7  
## 69 97.00000 267.0000 6.3 92 7 8  
## 70 97.00000 272.0000 5.7 92 7 9  
## 71 85.00000 175.0000 7.4 89 7 10  
## 72 42.12931 139.0000 8.6 82 7 11  
## 73 10.00000 264.0000 14.3 73 7 12  
## 74 27.00000 175.0000 14.9 81 7 13  
## 75 42.12931 291.0000 14.9 91 7 14  
## 76 7.00000 48.0000 14.3 80 7 15  
## 77 48.00000 260.0000 6.9 81 7 16  
## 78 35.00000 274.0000 10.3 82 7 17  
## 79 61.00000 285.0000 6.3 84 7 18  
## 80 79.00000 187.0000 5.1 87 7 19  
## 81 63.00000 220.0000 11.5 85 7 20  
## 82 16.00000 7.0000 6.9 74 7 21  
## 83 42.12931 258.0000 9.7 81 7 22  
## 84 42.12931 295.0000 11.5 82 7 23  
## 85 80.00000 294.0000 8.6 86 7 24  
## 86 108.00000 223.0000 8.0 85 7 25  
## 87 20.00000 81.0000 8.6 82 7 26  
## 88 52.00000 82.0000 12.0 86 7 27  
## 89 82.00000 213.0000 7.4 88 7 28  
## 90 50.00000 275.0000 7.4 86 7 29  
## 91 64.00000 253.0000 7.4 83 7 30  
## 92 59.00000 254.0000 9.2 81 7 31  
## 93 39.00000 83.0000 6.9 81 8 1  
## 94 9.00000 24.0000 13.8 81 8 2  
## 95 16.00000 77.0000 7.4 82 8 3  
## 96 78.00000 185.9315 6.9 86 8 4  
## 97 35.00000 185.9315 7.4 85 8 5  
## 98 66.00000 185.9315 4.6 87 8 6  
## 99 122.00000 255.0000 4.0 89 8 7  
## 100 89.00000 229.0000 10.3 90 8 8  
## 101 110.00000 207.0000 8.0 90 8 9  
## 102 42.12931 222.0000 8.6 92 8 10  
## 103 42.12931 137.0000 11.5 86 8 11  
## 104 44.00000 192.0000 11.5 86 8 12  
## 105 28.00000 273.0000 11.5 82 8 13  
## 106 65.00000 157.0000 9.7 80 8 14  
## 107 42.12931 64.0000 11.5 79 8 15  
## 108 22.00000 71.0000 10.3 77 8 16  
## 109 59.00000 51.0000 6.3 79 8 17  
## 110 23.00000 115.0000 7.4 76 8 18  
## 111 31.00000 244.0000 10.9 78 8 19  
## 112 44.00000 190.0000 10.3 78 8 20  
## 113 21.00000 259.0000 15.5 77 8 21  
## 114 9.00000 36.0000 14.3 72 8 22  
## 115 42.12931 255.0000 12.6 75 8 23  
## 116 45.00000 212.0000 9.7 79 8 24  
## 117 168.00000 238.0000 3.4 81 8 25  
## 118 73.00000 215.0000 8.0 86 8 26  
## 119 42.12931 153.0000 5.7 88 8 27  
## 120 76.00000 203.0000 9.7 97 8 28  
## 121 118.00000 225.0000 2.3 94 8 29  
## 122 84.00000 237.0000 6.3 96 8 30  
## 123 85.00000 188.0000 6.3 94 8 31  
## 124 96.00000 167.0000 6.9 91 9 1  
## 125 78.00000 197.0000 5.1 92 9 2  
## 126 73.00000 183.0000 2.8 93 9 3  
## 127 91.00000 189.0000 4.6 93 9 4  
## 128 47.00000 95.0000 7.4 87 9 5  
## 129 32.00000 92.0000 15.5 84 9 6  
## 130 20.00000 252.0000 10.9 80 9 7  
## 131 23.00000 220.0000 10.3 78 9 8  
## 132 21.00000 230.0000 10.9 75 9 9  
## 133 24.00000 259.0000 9.7 73 9 10  
## 134 44.00000 236.0000 14.9 81 9 11  
## 135 21.00000 259.0000 15.5 76 9 12  
## 136 28.00000 238.0000 6.3 77 9 13  
## 137 9.00000 24.0000 10.9 71 9 14  
## 138 13.00000 112.0000 11.5 71 9 15  
## 139 46.00000 237.0000 6.9 78 9 16  
## 140 18.00000 224.0000 13.8 67 9 17  
## 141 13.00000 27.0000 10.3 76 9 18  
## 142 24.00000 238.0000 10.3 68 9 19  
## 143 16.00000 201.0000 8.0 82 9 20  
## 144 13.00000 238.0000 12.6 64 9 21  
## 145 23.00000 14.0000 9.2 71 9 22  
## 146 36.00000 139.0000 10.3 81 9 23  
## 147 7.00000 49.0000 10.3 69 9 24  
## 148 14.00000 20.0000 16.6 63 9 25  
## 149 30.00000 193.0000 6.9 70 9 26  
## 150 42.12931 145.0000 13.2 77 9 27  
## 151 14.00000 191.0000 14.3 75 9 28  
## 152 18.00000 131.0000 8.0 76 9 29  
## 153 20.00000 223.0000 11.5 68 9 30

zxd<-data.frame(x<-c(NA,1,1,NA,2,1,5,3),  
 Y=c(4,2,NA,NA,7,2,1,8))  
zxd

## x....c.NA..1..1..NA..2..1..5..3. Y  
## 1 NA 4  
## 2 1 2  
## 3 1 NA  
## 4 NA NA  
## 5 2 7  
## 6 1 2  
## 7 5 1  
## 8 3 8

zxd2<-zxd[complete.cases(zxd),]  
zxd2

## x....c.NA..1..1..NA..2..1..5..3. Y  
## 2 1 2  
## 5 2 7  
## 6 1 2  
## 7 5 1  
## 8 3 8

zxd3<-na.aggregate(zxd)  
zxd3

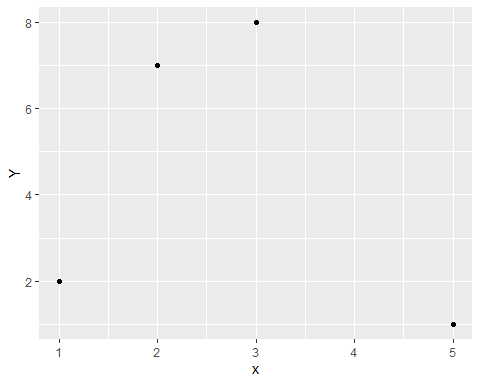
## x....c.NA..1..1..NA..2..1..5..3. Y  
## 1 2.166667 4  
## 2 1.000000 2  
## 3 1.000000 4  
## 4 2.166667 4  
## 5 2.000000 7  
## 6 1.000000 2  
## 7 5.000000 1  
## 8 3.000000 8

library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.6.3

ggplot(data=zxd,aes(x=x,y=Y))+  
 geom\_point()

## Warning: Removed 3 rows containing missing values (geom\_point).



ggplot(data=zxd3,aes(x=x,y=Y))+  
 geom\_point()

## Warning: Removed 2 rows containing missing values (geom\_point).

