novikovPrac4.R

Новиков Денис Владиславович ИНБО-05-22

2024-03-28

# Задание 1 ---------------------------------------------------------------  
  
# 1 -----------------------------------------------------------------------  
if (!require("readxl")) install.packages("readxl")

## Загрузка требуемого пакета: readxl

if (!require("lubridate")) install.packages("lubridate")

## Загрузка требуемого пакета: lubridate

##   
## Присоединяю пакет: 'lubridate'

## Следующие объекты скрыты от 'package:base':  
##   
## date, intersect, setdiff, union

library(readxl)  
library(lubridate)  
gas <- read\_excel('GAZ.xlsx',  
 col\_types = c('date', 'numeric', 'numeric', 'numeric',  
 'numeric', 'numeric', 'text', 'text', 'text'))  
colnames(gas) <- c("Date", "Preasure\_MPa", "Temp\_C", "Gas\_Prod\_m3\_per\_d",  
 "Cond\_m3\_per\_d", "Water\_m3\_per\_d", "ID", "Bush", "Group")  
  
# 2 -----------------------------------------------------------------------  
  
nrow(gas)

## [1] 28162

gas <- gas[complete.cases(gas), ]  
nrow(gas)

## [1] 28016

# 3 -----------------------------------------------------------------------  
  
colnames(gas)[colnames(gas) == "Temp\_C"] <- "Temp\_K"  
gas$Temp\_K <- gas$Temp\_K + 273.15  
head(gas)

## # A tibble: 6 × 9  
## Date Preasure\_MPa Temp\_K Gas\_Prod\_m3\_per\_d Cond\_m3\_per\_d  
## <dttm> <dbl> <dbl> <dbl> <dbl>  
## 1 2015-05-26 00:00:00 17.5 343. 361939 2.72  
## 2 2015-07-06 00:00:00 176 343. 349748 3.6   
## 3 2015-06-14 00:00:00 17.3 342. 310185 3.26  
## 4 2015-06-20 00:00:00 17.5 342. 265176 2.97  
## 5 2015-06-26 00:00:00 17.4 341. 263621 3   
## 6 2015-02-07 00:00:00 16 346. 446708 3.49  
## # ℹ 4 more variables: Water\_m3\_per\_d <dbl>, ID <chr>, Bush <chr>, Group <chr>

# 4 -----------------------------------------------------------------------  
  
gas$ID <- as.factor(gas$ID)  
gas$Bush <- as.factor(gas$Bush)  
gas$Group <- as.factor(gas$Group)  
str(gas)

## tibble [28,016 × 9] (S3: tbl\_df/tbl/data.frame)  
## $ Date : POSIXct[1:28016], format: "2015-05-26" "2015-07-06" ...  
## $ Preasure\_MPa : num [1:28016] 17.5 176 17.3 17.5 17.4 16 15.7 15.9 158 14.8 ...  
## $ Temp\_K : num [1:28016] 343 343 342 342 341 ...  
## $ Gas\_Prod\_m3\_per\_d: num [1:28016] 361939 349748 310185 265176 263621 ...  
## $ Cond\_m3\_per\_d : num [1:28016] 2.72 3.6 3.26 2.97 3 3.49 3.41 3.05 3.1 3.15 ...  
## $ Water\_m3\_per\_d : num [1:28016] 2.44 1.49 1.43 0.25 1.8 1.87 1.84 1.6 1.63 2 ...  
## $ ID : Factor w/ 81 levels "100","1008","1009",..: 58 58 58 58 58 58 58 58 58 58 ...  
## $ Bush : Factor w/ 13 levels "К-А","КП 1","КП 11",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Group : Factor w/ 4 levels "Западные Шады",..: 2 2 2 2 2 2 2 2 2 2 ...

# 5 -----------------------------------------------------------------------  
  
gas$Gas\_Cond\_Ratio <- gas$Gas\_Prod\_m3\_per\_d / gas$Cond\_m3\_per\_d  
gas$Gas\_Water\_Ratio <- gas$Gas\_Prod\_m3\_per\_d / gas$Water\_m3\_per\_d  
gas$Water\_Cond\_Ratio <- gas$Water\_m3\_per\_d / gas$Cond\_m3\_per\_d  
head(gas)

## # A tibble: 6 × 12  
## Date Preasure\_MPa Temp\_K Gas\_Prod\_m3\_per\_d Cond\_m3\_per\_d  
## <dttm> <dbl> <dbl> <dbl> <dbl>  
## 1 2015-05-26 00:00:00 17.5 343. 361939 2.72  
## 2 2015-07-06 00:00:00 176 343. 349748 3.6   
## 3 2015-06-14 00:00:00 17.3 342. 310185 3.26  
## 4 2015-06-20 00:00:00 17.5 342. 265176 2.97  
## 5 2015-06-26 00:00:00 17.4 341. 263621 3   
## 6 2015-02-07 00:00:00 16 346. 446708 3.49  
## # ℹ 7 more variables: Water\_m3\_per\_d <dbl>, ID <fct>, Bush <fct>, Group <fct>,  
## # Gas\_Cond\_Ratio <dbl>, Gas\_Water\_Ratio <dbl>, Water\_Cond\_Ratio <dbl>

# 6 -----------------------------------------------------------------------  
  
gas\_2018 <- subset(gas, year(Date) == 2018)  
nrow(gas\_2018)

## [1] 4669

# 7 -----------------------------------------------------------------------  
  
gas\_111 = gas[gas$ID == 111, ]  
nrow(gas\_111)

## [1] 136

# 8 -----------------------------------------------------------------------  
  
k\_more\_than\_2 <- table(subset(gas, Water\_m3\_per\_d > 2)$ID)  
ID\_less\_than\_2 <- names(k\_more\_than\_2[k\_more\_than\_2 == 0])  
ID\_less\_than\_2

## [1] "1010" "1024" "103" "107" "109" "118" "165" "7"

length(ID\_less\_than\_2)

## [1] 8

# 9 -----------------------------------------------------------------------  
  
k\_less\_than\_1000 <- table(subset(gas, Gas\_Prod\_m3\_per\_d < 1000)$ID)  
ID\_more\_than\_1000 <- names(k\_less\_than\_1000[k\_less\_than\_1000 == 0])  
ID\_more\_than\_1000

## [1] "1008" "101" "1010" "1011" "1012" "1013" "1014" "1015" "1016" "1017"  
## [11] "1018" "1019" "102" "1020" "1021" "1022" "1023" "1024" "1025" "1026"  
## [21] "1027" "103" "1030" "1031" "1034" "1036" "1037" "1038" "1039" "104"   
## [31] "1040" "1041" "1043" "1044" "1045" "1047" "1048" "1049" "105" "1051"  
## [41] "1052" "1053" "1054" "1055" "1056" "1057" "1058" "106" "1081" "109"   
## [51] "111" "114" "160" "161" "162" "163" "164" "165" "171" "173G"  
## [61] "17B" "18B" "20B" "242" "301" "302" "305" "306" "7"

length(ID\_more\_than\_1000)

## [1] 69

max <- 12500  
k\_less\_than\_1000 <- table(subset(gas, Gas\_Prod\_m3\_per\_d < max)$Bush)  
Bush\_more\_than\_1000 <- names(k\_less\_than\_1000[k\_less\_than\_1000 == 0])  
Bush\_more\_than\_1000

## [1] "КП 11" "КП 12" "КП 3"

max

## [1] 12500

# 10 ----------------------------------------------------------------------  
  
gas\_2018 <- subset(gas, year(Date) == 2018)  
agg\_data <- aggregate(Gas\_Prod\_m3\_per\_d ~ Group, data = gas\_2018, sum)  
colnames(agg\_data)[2] <- "Total\_Prod\_2018"  
agg\_data

## Group Total\_Prod\_2018  
## 1 Западные Шады 70176357  
## 2 Кувачи-Алат 85671253  
## 3 Северные Шады 55608131  
## 4 Хаузак 308613180

max\_group <- agg\_data[which.max(agg\_data$Total\_Prod\_2018), 1]  
max\_group

## [1] Хаузак  
## Levels: Западные Шады Кувачи-Алат Северные Шады Хаузак

# 11 ----------------------------------------------------------------------  
  
gas\_2018 <- subset(gas, year(Date) == 2018)  
agg\_data <- aggregate(Water\_m3\_per\_d ~ Bush, data = gas\_2018, sum)  
colnames(agg\_data)[2] <- "Total\_Prod\_2018"  
agg\_data

## Bush Total\_Prod\_2018  
## 1 К-А 284.788  
## 2 КП 1 1494.829  
## 3 КП 11 319.241  
## 4 КП 12 389.726  
## 5 КП 14 180.844  
## 6 КП 16 490.239  
## 7 КП 19 374.107  
## 8 КП 2 632.416  
## 9 КП 242 191.167  
## 10 КП 3 114.592  
## 11 КП 301 126.208  
## 12 КП 9 256.456  
## 13 СШ 585.523

max\_bush <- agg\_data[which.max(agg\_data$Total\_Prod\_2018), 1]  
max\_bush

## [1] КП 1  
## 13 Levels: К-А КП 1 КП 11 КП 12 КП 14 КП 16 КП 19 КП 2 КП 242 КП 3 ... СШ

# 12 ----------------------------------------------------------------------  
  
gas\_data <- gas[is.finite(gas$Gas\_Water\_Ratio), ]  
agg\_data <- aggregate(Gas\_Water\_Ratio ~ Bush, data = gas\_data, mean)  
colnames(agg\_data)[2] <- "Mean\_Gas\_Water\_Ratio"  
agg\_data

## Bush Mean\_Gas\_Water\_Ratio  
## 1 К-А 1100133.1  
## 2 КП 1 1901441.3  
## 3 КП 11 463532.7  
## 4 КП 12 267186.6  
## 5 КП 14 870718.5  
## 6 КП 16 334753.3  
## 7 КП 19 416871.1  
## 8 КП 2 281412.9  
## 9 КП 242 5450558.8  
## 10 КП 3 178675.8  
## 11 КП 301 1483464.5  
## 12 КП 9 321463.5  
## 13 СШ 298763.7

max\_mean\_bush <- agg\_data[which.max(agg\_data$Mean\_Gas\_Water\_Ratio), 1]  
max\_mean\_bush

## [1] КП 242  
## 13 Levels: К-А КП 1 КП 11 КП 12 КП 14 КП 16 КП 19 КП 2 КП 242 КП 3 ... СШ