Голоскокоа Артем, РИ-411055

РАБОТА С ДАННЫМИ

names(data.df) <- c("year", "month", "day", seq(0,23))

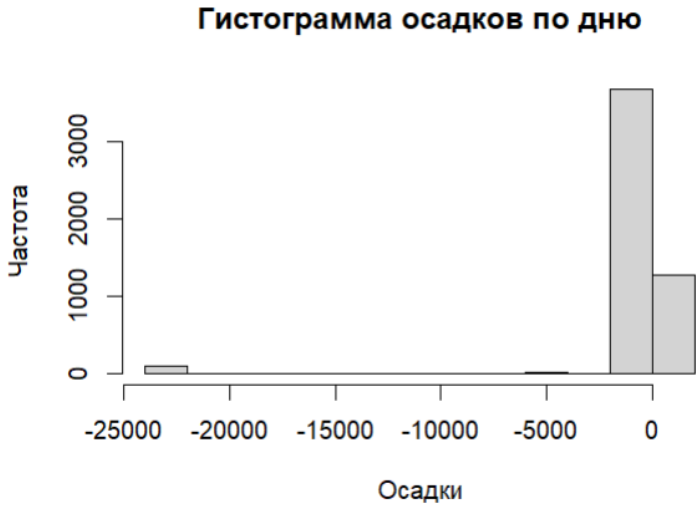
Эта строка присваивает названия колонкам в data.df, где первые три колонки будут называться year, month, day, а оставшиеся 24 колонки будут иметь названия, соответствующие часам в сутках: от 0 до 23.

data.df$daily <- rowSums(data.df[, 4:27])

Добавление новой колонки с названием daily, в которую записывается сумма крайних правых 24 колонок.

hist(data.df$daily, main="Гистограмма осадков по дню", xlab="Осадки", ylab="Частота")

Построение гистограммы по колонке daily



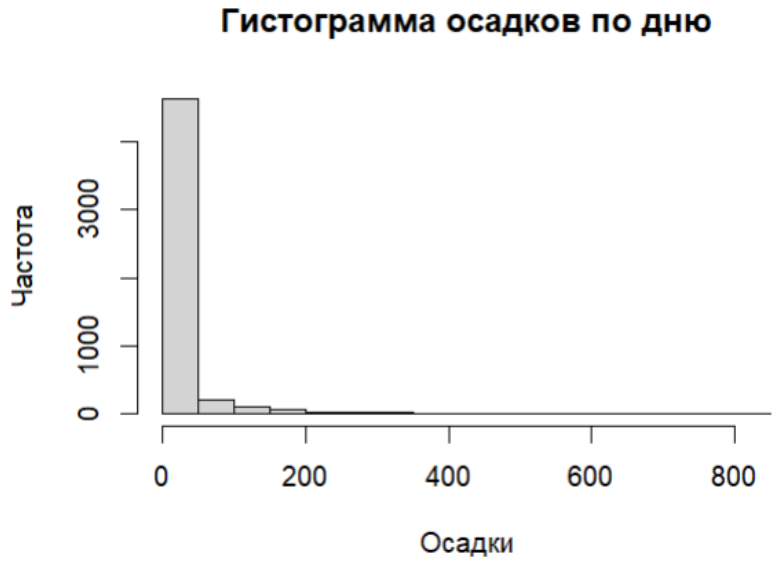
Вывод: в наборе данных есть отрицательные значения количества осадков

Создаем датафрейм fixed.df и заменяем все отрицательные элементы на 0:

fixed.df[, 4:27][data.df[, 4:27] < 0] <- 0

Построение гистограммы по колонке daily:

hist(fixed.df$daily, main="Гистограмма осадков по дню", xlab="Осадки", ylab="Частота")



Вывод: Построенная гистограмма корректнее, благодаря замене отрицательных значений в датасете на 0

СИНТАКСИС И ТИПИЗИРОВАНИЕ

v <- c("4", "8", "15", "16", "23", "42")

Результат: создаётся вектор v, который содержит строки (символы) "4", "8", "15", "16", "23", "42". Эти значения интерпретируются как текстовые данные (тип данных — character), а не как числа.

max(v)

Результат команды: "8", т.к. значения интерпретируются как текстовые

sort(v)

Результат: функция sort() отсортирует элементы вектора как строки

[1] "15" "16" "23" "4" "42" "8"

sum(v)

Ошибка: функция sum() ожидает числовой вектор, чтобы вычислить сумму, но в данном случае вектор состоит из строк, поэтому R выдаст ошибку

v2 <- c("5", 7, 12)

Результат: создаётся вектор v2, который содержит элементы "5", 7 и 12. "5" — это строка (тип данных character), а 7 и 12 — числовые значения (тип данных numeric).

v2[2] + 2[3]

Ошибка:

Индексация вектора v2 осуществляется как v2[n], где n — это индекс. Правильная команда: v2[2] + v2[3]

df3 <- data.frame(z1="5", z2=7, z3=12)

Результат: создаётся датафрейм df3, который содержит одну строку и три колонки

z1: "5" (строка),
z2: 7 (число),
z3: 12 (число)

df3[1,2] + df3[1,3]

Результат: 19 (Сложение значений из первой строки, второй и третьей колонок)

l4 <- list(z1="6", z2=42, z3="49", z4=126)

Результат: создаётся список l4 с элементами
z1: "6" (строка),
z2: 42 (число),
z3: "49" (строка),
z4: 126 (число)

l4[[2]] + l4[[4]]

Результат: 168

l4[2] + l4[4]

Ошибка:

В этой команде происходит ошибка, потому что l4[2] и l4[4] возвращают списки, а не сами элементы. В R, обращение через одинарные квадратные скобки [] для списков возвращает подсписки, а не значения. Для арифметических операций нужно использовать двойные скобки [[ ]], которые возвращают сами элементы списка.

РАБОТА С ФУНКЦИЯМИ И ОПЕРАТОРАМИ

Числа от 1 до 10000 с инкрементом 372:
seq(1, 10000, by=372)
1 — начальное значение.
10000 — конечное значение
by=372 — шаг, с которым будут создаваться числа.

Числа от 1 до 10000 длиной 50:
seq(1, 10000, length.out=50)
1 — начальное значение.
10000 — конечное значение.
length.out=50 — количество чисел, которые должно быть в последовательности.

rep(1:5, times=3):
Эта команда повторяет весь вектор 1:5 три раза подряд. Результат будет:
[1] 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5

rep(1:5, each=3):
Эта команда повторяет каждый элемент вектора 1:5 по три раза.
[1] 1 1 1 2 2 2 3 3 3 4 4 4 5 5 5