**РАБОТА С ДАННЫМИ**

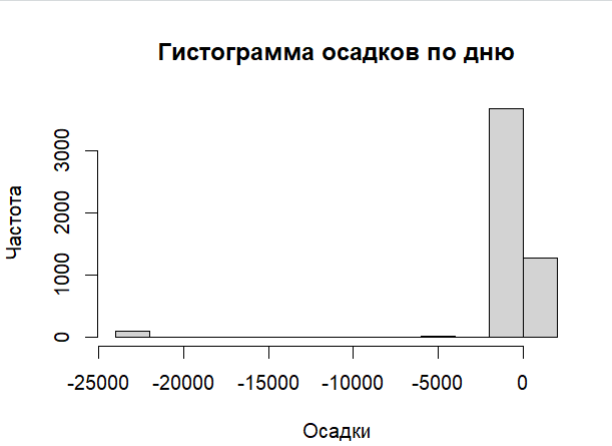
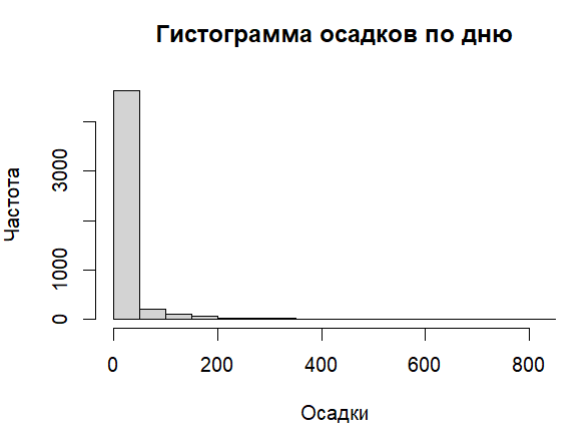


Рисунок 1 - Построение гистограммы по колонке daily

Вывод: видим что гистограмма некорректна и в наборе данных есть отрицательные значения количества осадок



Вывод: гистограмма получилась корректнее, но из-за того, что мы не знаем правильные значения ячеек, где были отрицательные значения, она всё равно не идеально точная

**СИНТАКСИС И ТИПИЗИРОВАНИЕ**

Работа с вектором v:

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

Здесь создаётся вектор v из строковых элементов (все элементы заключены в кавычки, то есть это строки, а не числа).

max(v)

Команда max() ищет максимальное значение в векторе, однако работает она корректно только с числовыми значениями. Поскольку все элементы вектора — строки, функция max() сравнивает их как строки по лексикографическому порядку, а не по числовому значению.

Результат: Вернётся "8", так как в лексикографическом порядке "8" больше, чем "4", "15" и остальные строки.

sort(v)

Команда sort() сортирует элементы вектора. Поскольку элементы вектора — строки, они будут отсортированы в алфавитном (лексикографическом) порядке.

Результат: Вернётся отсортированный вектор: ("15", "16", "23", "4", "42", "8"). Числа в строковом формате сортируются по первым символам.

sum(v)

Функция sum() ожидает числовой вектор для сложения его элементов. Поскольку вектор v содержит строки, эта операция приведёт к ошибке.

Результат: Ошибка, потому что невозможно выполнить операцию сложения строковых значений.

Работа с вектором v2:

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

Здесь создаётся вектор v2, состоящий из элементов "5" (строка), 7 и 12 (числа).

v2[2] + 2[3]

Выражение пытается сложить второй элемент вектора v2 и третий элемент числа 2. Здесь присутствует синтаксическая ошибка, потому что запись 2[3] не имеет смысла в R. Индексация чисел не поддерживается таким образом.

Результат: Ошибка синтаксиса, потому что выражение 2[3] недопустимо.

Работа с data.frame:

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

Создаётся датафрейм df3 с колонками z1, z2 и z3. Столбец z1 содержит строку "5", а столбцы z2 и z3 — числа 7 и 12.

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

Здесь выполняется сложение значения из первой строки и второго столбца (значение 7) со значением из первой строки и третьего столбца (значение 12).

Результат: Вернётся 19, потому что 7 + 12 = 19.

Работа со списком l4:

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

Создаётся список l4, где z1 и z3 — строки, а z2 и z4 — числа.

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

Здесь идёт обращение к элементам списка по их порядковым номерам, где l4[[2]] — это число 42, а l4[[4]] — это число 126. Оба элемента — числовые, поэтому сложение корректно.

Результат: Вернётся 168, так как 42 + 126 = 168.

l4[2] + l4[4]

В данном случае обращение к элементам списка через l4[2] и l4[4] возвращает подсписки, а не сами значения. Операция сложения подсписков не определена в R.

Результат: Ошибка, так как операция сложения подсписков невозможна.

**РАБОТА С ФУНКЦИЯМИ И ОПЕРАТОРАМИ**

Для создания последовательности с шагом 372 используют функцию seq() с параметром by.

seq(1, 10000, by = 372)

1 — начальное значение.

10000 — конечное значение.

by = 372 — шаг последовательности

Числа от 1 до 10000 длиной 50:

Для создания последовательности определённой длины используют параметр length.out.

seq(1, 10000, length.out = 50)

1 — начальное значение.

10000 — конечное значение.

length.out = 50 — длина последовательности.

Различия:

times повторяет весь вектор целиком указанное число раз.

each повторяет каждый элемент вектора заданное количество раз