Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение

высшего образования

**«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»**

**(ФГБОУ ВО «КубГУ»)**

**Факультет компьютерных технологий и прикладной математики**

**Кафедра вычислительных технологий**

**Отчёт лабораторная №3**

**Дисциплина: Обработка больших данных**

Работу выполнил \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Масенко М.С.

Направление подготовки 02.03.02 Фундаментальная информатика и

информационные технологии

Преподаватель \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Яхонтов А.А.

Краснодар

2025

**Тема:** Работа с графиками и диаграммами в R.

**Цель**: научиться работать с графиками.

**Задание:**

1. Собрать данные согласно варианту по указанному виду спорта в заданных странах.
2. Вывести графики динамики олимпийских достижений заданной страны по виду спорта относительно временной шкалы: источник (https://olympteka.ru/olymp/sport/country\_medalists/23/ger.html) столбчатую диаграмму по количеству мест 1-8 (спортсменов заданных стран) по каждой Олимпиаде по назначенному виду спорта, круговую диаграмму по количеству первых мест в каждой из олимпиад, функциональные графики - тенденции изменения количества призовых мест отдельно по мужчинам и женщинам за последние 30 лет.
3. Вывести графики изменения спортивных достижений 1) по золотым медалям и 2) по призовым 3-местам по 7-и странам-призерам (разными цветами и точками) за последние 6 олимпиад. http://olympteka.ru/olymp/game/medals/51.html, например,
4. Используя тот же источник, выбрать информацию за последние 6 олимпиад по заданному виду спорта, аккумулировать данные по мужчинам и женщинам (отобразить на одном графике разными линиями), вывести динамику и статистику на графики (функциональный, столбчатый, круговой), используя изученные приемы оформления, в том числе массив графиков на одном поле.

Вариант: 26.

Страна: Щвеция

Вид спорта: лыжные гонки.

**Ход работы**:

Соберем статистику достижений Щвеции по лыжным гонкам за последние 30 лет. Все данные перенесем в таблицы excel, разделив мужчин и женщин. Лыжные гонки проводятся только во время зимних олимпийских игр.

Изображение выглядит как текст, число

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 1 – Датасет «Мужчины»

Изображение выглядит как текст, число

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 2 – Датасет «Женщины»

Необходимо построить столбчатые диаграммы по количеству мест 1-8 по каждой олимпиаде по лыжным гонкам. Слева выводим статистику по женщинам, справа – по мужчинам.

Изображение выглядит как диаграмма, График, снимок экрана, Шрифт

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 3 – Столбчатые диаграммы количества мест 1-8 по категориям

По построенным диаграммам можно сделать выводы о том, что женщины чаще занимают призовые места, а мужчины в общем занимают больше мест.

Далее построим круговые диаграммы по количеству 1-х мест по каждой олимпиаде для женщин и мужчин.

Изображение выглядит как диаграмма, снимок экрана, круг, Красочность

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 4 – Круговые диаграммы количества 1-х мест

По этим диаграммам видно, что за всё время женщины получили больше золотых медалей, чем мужчины. Последняя на текущий момент золотая медаль завоёвана в 2022 году женщинами.

Выведем функциональные графики, которые отображают тенденции изменения количества призовых мест по категориям за последние 30 лет.

Изображение выглядит как линия, диаграмма, текст, График

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 5 – Тенденции изменения количества призовых мест за последние 30 лет по категориям

По графикам видно, что в период с 1992 по 2014 годы наблюдался рост кол-ва призовых мест у мужчин, однако последние два года мужчины не получали призовые места. Женщины в период с 1992 по 2002 годы не занимали призовых мест, однако с 2006 года наблюдает рост кол-ва полученных призовых мест с небольшим спадом в 2018-2022 годах.

Создадим график изменения спортивных достижений по золотым медалям по 7-ми странам призерам за последние 6 Олимпиад, предварительно собрав данные.

Изображение выглядит как текст, Шрифт, число, чек

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 6 – Датасет для выявления динамики количества 1-х мест за последние 6 олимпиад по 7 странам-призерам

Изображение выглядит как линия, График, диаграмма

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 7 – Динамика количества 1-х мест за последние 6 олимпиад по 7 странам-призерам

Можно заметить, что США одерживали больше всего 1-ых мест. Нидерланды показывают самую стабильную статистику по полученным 1-ым местам во всех годах.

Теперь выведем график изменения спортивных достижений по призовым местам по тем же странам за последние 6 олимпиад. Занесём данные в таблицу:

Изображение выглядит как текст, число, Шрифт

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 8 – Датасет для выявления динамики количества призовых мест за последние 6 олимпиад по 7 странам-призерам

Изображение выглядит как линия, График, диаграмма, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 9 – Динамика количества призовых мест за последние 6 олимпиад по 7 странам-призерам

В целом, график не сильно изменился: больше всего побед одержали США, однако Нидерланды показывают немного более разнообразную статистику. Также можно заметить, что в годы, когда США выигрывают наибольшее кол-во медалей, общая тенденция по занятым призовым местам у остальных стран также растет.

Взяв информацию о призовых местах Щвеции по лыжным гонкам за последние 6 олимпиад, аккумулируем данные по женщинам и мужчинам и выведем динамику и статистику на графики.

Изображение выглядит как диаграмма, График, линия, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

Рисунок 10 – Динамика и статистика по призовым местам за последние 6 олимпиад по фигурному катанию

По данным на графиках видно, что кол-во занятых призовых мест на последних 6 олимпиадах одинаково, однако, наблюдается разная тенденция получения призовых мест. У мужчин наблюдается рост получения призовых мест с 1992 по 2014, далее спад. У женщин спад с 1992 по 2002, далее высокий рост.

**Вывод:** проделав лабораторную работу, были изучены способы графического представления данных и реализованы на практике.

**Листинг программы:**

library(openxlsx)

quartz(width = 10, height = 5)

men <- read.xlsx("/Users/mariamasenko/University/6\_semestr/big data/lab3/Лыжные гонки.xlsx", sheet = 1)

women <- read.xlsx("/Users/mariamasenko/University/6\_semestr/big data/lab3/Лыжные гонки.xlsx", sheet = 2)

# Столбчатые диаграммы

bc\_men\_count <- apply(men[,-1], 2, sum)

bc\_women\_count <- apply(women[,-1], 2, sum)

max\_val <- max(bc\_men\_count)

par(mfrow = c(1, 2))

barplot(bc\_men\_count,

names.arg = colnames(men)[-1],

col = "#59ffbf",

main = "Мужчины",

xlab = "Места",

ylab = "Количество",

ylim = c(0, max\_val+1),

)

barplot(bc\_women\_count,

names.arg = colnames(women)[-1],

col = "#ae43ff",

main = "Женщины",

xlab = "Места",

ylab = "Количество",

ylim = c(0, max\_val+1),

)

# года, в которых были заняты первые места

years\_men <- men$Год[men$`1` > 0]

years\_women <- women$Год[women$`1` > 0]

# кол-во первых мест

first\_place\_men <- sum(men$`1` > 0)

first\_place\_women <- sum(women$`1` > 0)

# круговые диаграммы

par(mfrow = c(1, 2))

legend\_colors <- c("darkgreen", "blue", "red", "orange", "yellow", "green", "purple", "lightgreen")

# Для мужчин

if (length(years\_men) == 0) {

pie(c(1, 0), labels = "0", main = "Мужчины", col = "gray")

legend("topright", legend = c("Нет первых мест"), fill = "gray")

} else {

men\_table <- table(years\_men) # считаем количество первых мест по годам

pie(as.numeric(men\_table),

labels = as.numeric(men\_table), # Отображаем количество призовых мест на диаграмме

main = "Мужчины",

col = legend\_colors[1:length(men\_table)],

cex = 0.6)

legend("topright",

legend = names(men\_table), # Отображаем года в легенде

fill = legend\_colors[1:length(men\_table)],

cex = 0.6)

}

# Для женщин

if (length(years\_women) == 0) {

pie(c(1, 0), labels = "0", main = "Женщины", col = "gray")

legend("topright", legend = c("Нет первых мест"), fill = "gray")

} else {

women\_table <- table(years\_women)

pie(as.numeric(women\_table),

labels = as.numeric(women\_table), # Отображаем количество призовых мест на диаграмме

main = "Женщины",

col = legend\_colors[1:length(women\_table)],

cex = 0.6)

legend("topright",

legend = names(women\_table), # Отображаем года в легенде

fill = legend\_colors[1:length(women\_table)],

cex = 0.6)

}

# функциональные графики

par(mfrow = c(1, 2))

plot\_graph <- function(df, color, title) {

total\_places <- rowSums(df[, 2:4], na.rm = TRUE)

matplot(df$Год, total\_places,main = title, type = "b", pch = 19, lty = 1, col = color,

xlab = "Год", ylab = "Количество мест", xaxt = "n",

ylim = c(0, max(df[, 2]\*4)))

axis(1, at = df$Год, labels = df$Год)

}

plot\_graph(men, "blue", "Тенденции изменения кол-ва призовых мест\n для мужчин за последние 30 лет")

plot\_graph(women, "#ae43ff", "Тенденции изменения кол-ва призовых мест\n для женщин за последние 30 лет")

# график изменения спортивных достижений (золото)

gold <- read.xlsx("/Users/mariamasenko/University/6\_semestr/big data/lab3/Лыжные гонки.xlsx", sheet = 3)

par(mar = c(6, 4, 4, 11))

matplot(gold$Год,

gold[, -1],

main = "Изменение кол-ва 1-х мест по 7-ми странам-призерам\n за последние 6 олимпиад",

type = "b", pch = 19,

lty = 1,

lwd = 2,

col = legend\_colors,

xlab = "Год", ylab = "Количество мест (золото)",

xaxt = "n",

ylim = c(0, max(gold[, -1]+10)))

axis(1, at = gold$Год, labels = gold$Год)

legend("topright", legend = colnames(gold)[-1], xpd = TRUE, bty = "n", inset = -c(0.25, 0), fill = legend\_colors)

# график изменения спортивных достижений (все призовые)

prize\_places <- read.xlsx("/Users/mariamasenko/University/6\_semestr/big data/lab3/Лыжные гонки.xlsx", sheet = 4)

par(mar = c(5, 4, 4, 11))

matplot(prize\_places$Год,

prize\_places[, -1],

main = "Изменение кол-ва призовых мест по 7-ми странам-призерам\n за последние 6 олимпиад",

type = "b", pch = 19,

lty = 1,

lwd = 2,

col = legend\_colors,

xlab = "Год", ylab = "Количество мест (призовые)",

xaxt = "n",

ylim = c(0, max(prize\_places[, -1]+10)))

axis(1, at = prize\_places$Год, labels = prize\_places$Год)

legend("topright", legend = colnames(prize\_places)[-1], xpd = TRUE, bty = "n", inset = -c(0.25, 0), fill = legend\_colors)

# динамика лыжных гонок по женщинам и мужчинам за последние 6 Олимпиад

all <- cbind(women[,1:4], men[,2:4])

total\_places1 <- rowSums(women[, 2:4], na.rm = TRUE)

total\_places2 <- rowSums(men[, 2:4], na.rm = TRUE)

total\_places\_all <- cbind(women[, 1], total\_places1, total\_places2)

year <- all[1:6, ]

par(mfrow = c(1, 3), oma = c(0, 0, 2, 0))

matplot(total\_places\_all[, 1],

total\_places\_all,

type = "b", pch = 19,

lty = c(1, 3),

col = c("blue", "darkgreen"),

xlab = "Год", ylab = "Количество мест (призовые)",

xaxt = "n",

lwd = 3,

ylim = c(0, max(year[, -1])+5))

axis(1, at = year$Год, labels = year$Год)

legend('topleft', c('Мужчины', 'Женщины'), lty = c(1, 3), cex = 0.8, lwd=4, col = c("blue", "darkgreen"))

colnames(total\_places\_all) <- c("Год", "Женщины", "Мужчины")

total\_places\_all\_t <- t(total\_places\_all)

total\_places\_all\_t\_sorted <- total\_places\_all\_t[, order(total\_places\_all\_t[1, ])]

barplot(as.matrix(total\_places\_all\_t\_sorted[-1,]),

beside = TRUE,

col = c("darkgreen", "blue"),

xlab = "Год",

names.arg = total\_places\_all\_t\_sorted[1,],

ylab = "Количество мест (призовые)",

ylim = c(0,8),

cex.names = 0.8)

legend('topright', c('Женщины', 'Мужчины'), fill=c("darkgreen", "blue"))

total\_places\_pie <- apply(total\_places\_all[, -1], 2, sum)

pie(total\_places\_pie,

labels = c(total\_places\_pie[1], total\_places\_pie[2]),

col=c("darkgreen", "blue"),

sub="Кол-во призовых мест за 6 олимпиад\n (женщины, мужчины)")

legend('topright', c('Женщины', 'Мужчины'), fill=c("darkgreen", "blue"))

mtext("Общее количество призовых мест на 6 олимпиадах по лыжным гонкам",

outer = TRUE, cex = 1, font = 2)