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 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Гаранина Л.В.

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: 22.

Страна: Германия

Вид спорта: фигурное катание.

**Ход работы**:

Соберем статистику достижений Германии по фигурному катанию за последние 30 лет.



Таблица 1 – Статистика по фигурному катанию Германии

Все данные занесем в таблицы эксель, разделяя мужчин, женщин и спортивные пары. Фигурное катание проводится только в период зимних Олимпийских игр.

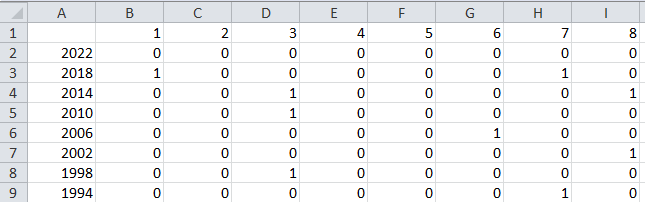
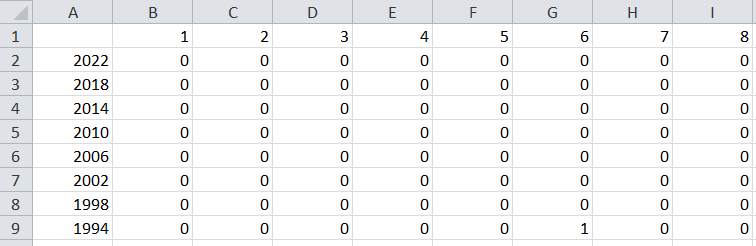


Рисунок 2 – Датасет "микст"

Рисунок 3 – Датасет "Женщины"

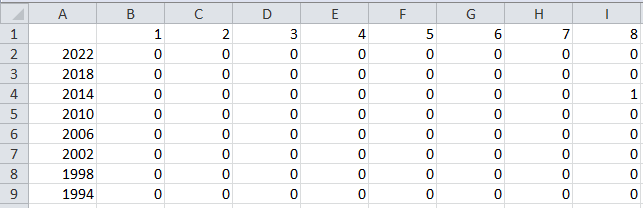


Рисунок 4 – Датасет "Мужчины"

Необходимо построить столбчатые диаграммы по количеству мест 1-8 по каждой Олимпиаде по фигурному катанию. Слева выводим статистику по женщинам, справа – по мужчинам, в центре – по миксу, то есть групповым-парным выступлениям.

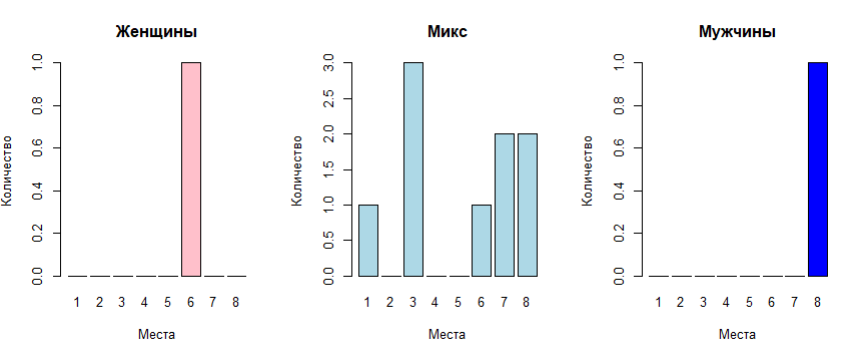


Рисунок 5 – Столбчатые диаграммы количества мест 1-8 по категориям

По построенным диаграммам можно сказать, что женщины и мужчины всего по одному разу заняли место из топ 8 за последние 30 лет, групповые команды в этом преуспели больше.

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

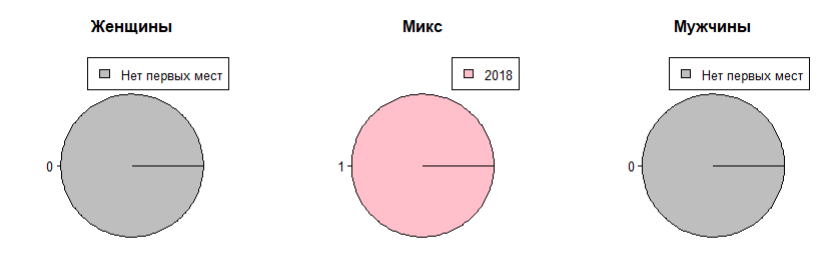


Рисунок 6 – Круговые диаграммы количества 1-х мест

Исходя из диаграмм, видно, что женщины и мужчины ни разу не выигрывали золотые медали. Единственный раз первое место занял микс в 2018 году. Это подчеркивает сильную зависимость результатов от команды в смешанных категориях, в отличие от традиционных соревнований среди женщин и мужчин, где за все время не было зафиксировано победных позиций.

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

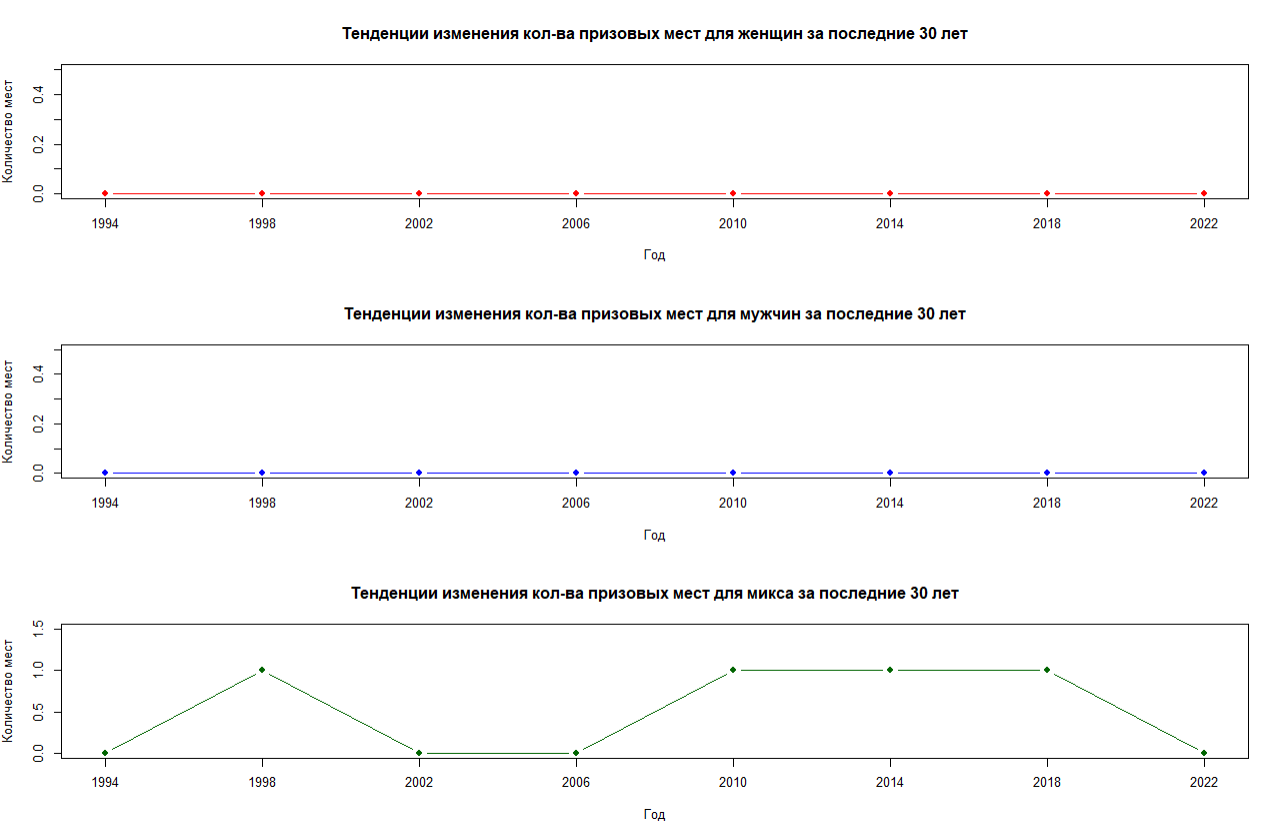


Рисунок 7 – Тенденции изменения количества призовых мест за последние 30 лет по категориям

По графикам видно, что женщины и мужчины стабильно не занимали призовые места. Все призовые позиции в основном были заняты смешанными командами: с 2010 года по 2018 микс зарабатывал по одному призовому месту – это самая длинная серия наград.

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

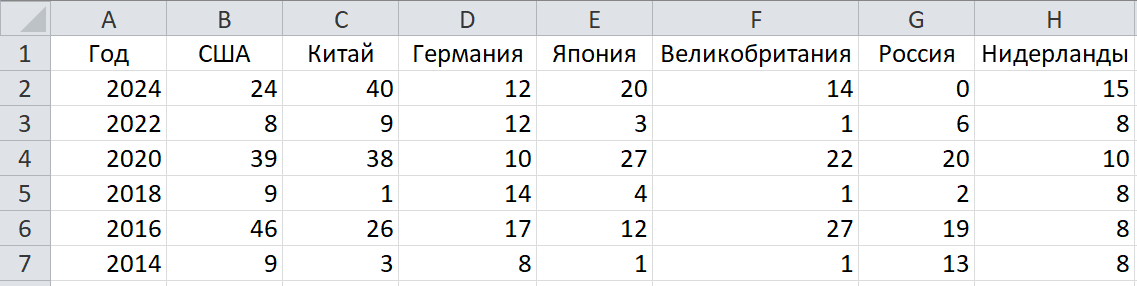


Рисунок 8 – Датасет для выявления динамики количества 1-х мест за последние 6 олимпиад по 7 странам-призерам

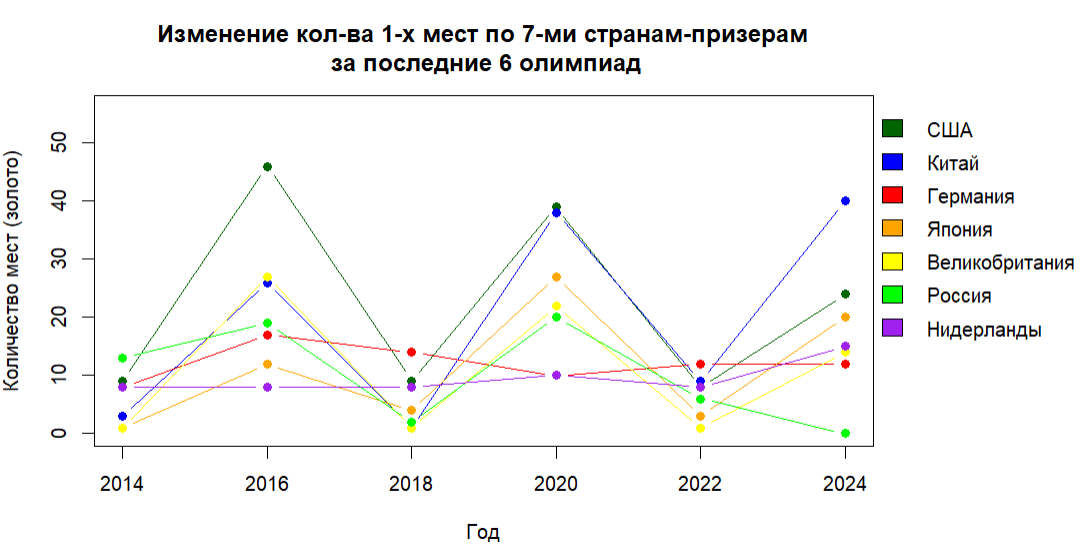


Рисунок 9 – Динамика количества 1-х мест за последние 6 олимпиад по 7 странам-призерам

В основном страны зарабатывали больше всего медалей во время летних Олимпийских игр, где в основном доминировали США и Китай. Нидерланды и Германия вели себя стабильно ровно и во время зимних Олимпиад.

Далее сделаем то же самое, но для всех призовых мест.

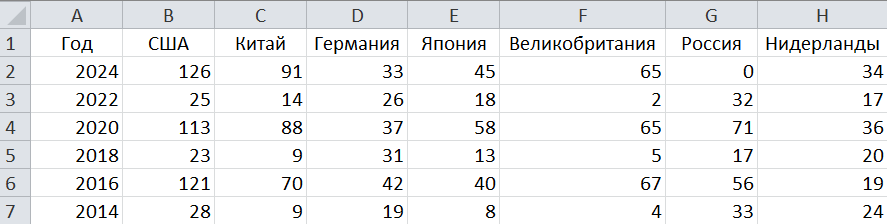


Рисунок 10 – Датасет для выявления динамики количества призовых мест за последние 6 олимпиад по 7 странам-призерам

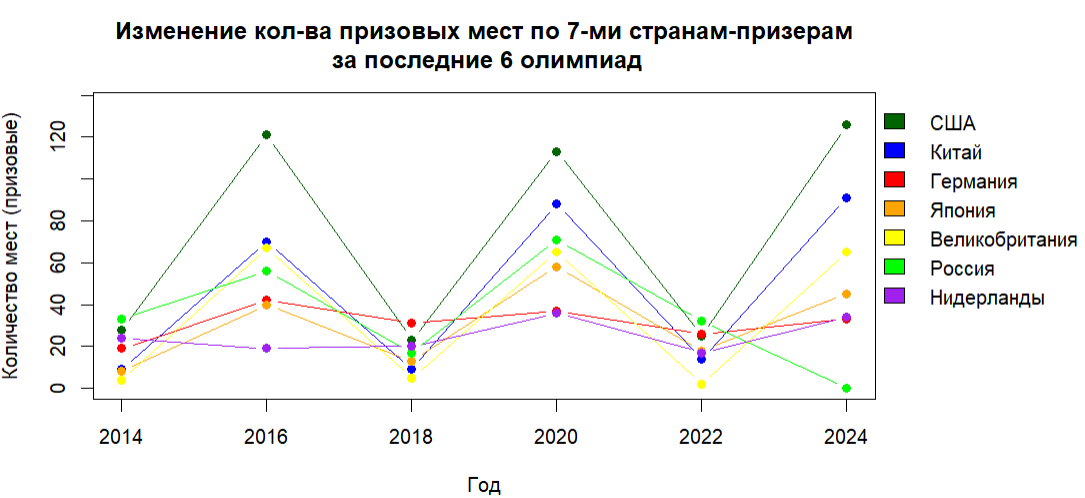


Рисунок 11 – Динамика количества призовых мест за последние 6 олимпиад по 7 странам-призерам

В целом, график не сильно изменился: в основном страны зарабатывали больше всего медалей на летних Олимпийских играх, где США сильно выделяется на фоне побед.

Взяв информацию о призовых местах Германии по фигурному катанию за последние 6 Олимпиад, аккумулируем данные по миксу, женщинам и мужчинам и выведем динамику и статистику на графики.

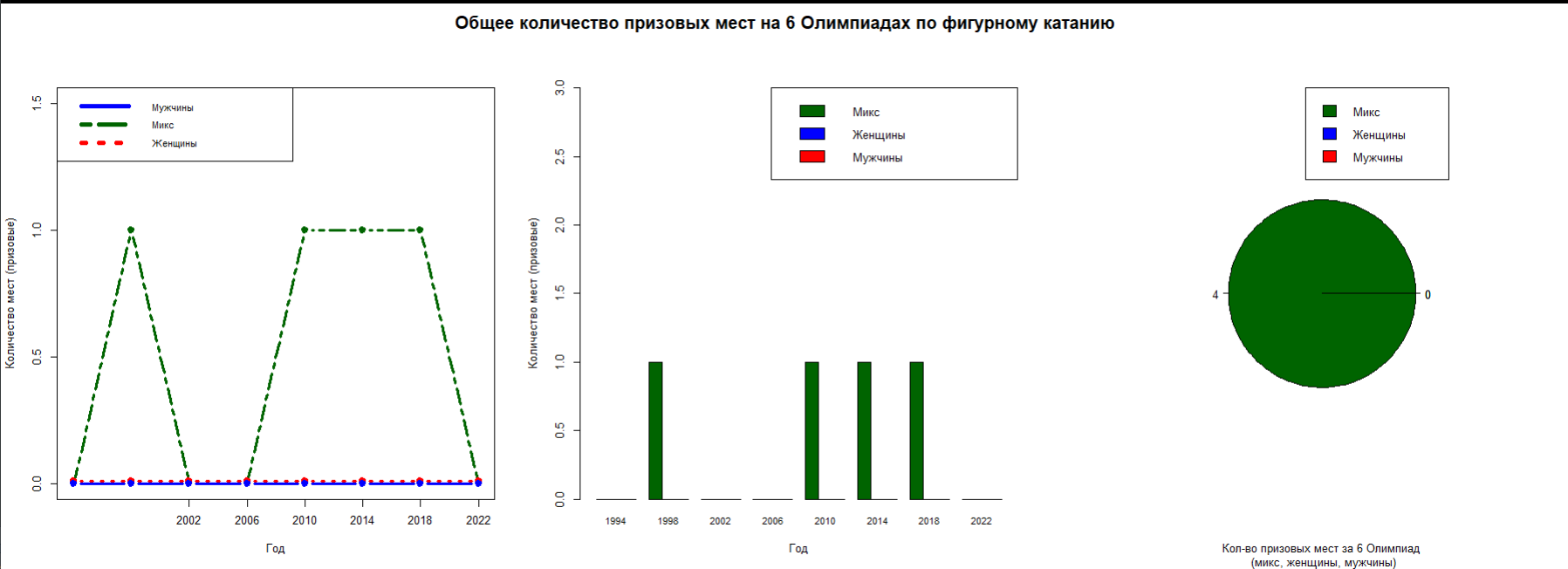


Рисунок 12 – Динамика и статистика по призовым местам за последние 6 Олимпиад по фигурному катанию

По данным графиков можно сказать, что у Германии в фигурном катании лучше всего получается выигрывать в командных соревнованиях, чем в индивидуальных.

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

**Листинг программы:**

library(readxl)

setwd("C:/Users/Людмила/Desktop/big data/rstudioLAB")

df\_mix <- read\_excel("Olimp.xlsx", sheet = 1)

df\_woman <- read\_excel("Olimp.xlsx", sheet = 2)

df\_man <- read\_excel("Olimp.xlsx", sheet = 3)

# Столбчатые диаграммы

df\_mix\_count <- apply(df\_mix[,-1], 2, sum)

df\_woman\_count <- apply(df\_woman[,-1], 2, sum)

df\_man\_count <- apply(df\_man[,-1], 2, sum)

par(mfrow = c(1, 3))

barplot(df\_woman\_count,

names.arg = colnames(df\_woman)[-1],

col = "pink",

main = "Женщины",

xlab = "Места",

ylab = "Количество",

)

barplot(df\_mix\_count,

names.arg = colnames(df\_mix)[-1],

col = "lightblue",

main = "Микс",

xlab = "Места",

ylab = "Количество")

barplot(df\_man\_count,

names.arg = colnames(df\_man)[-1],

col = "blue",

main = "Мужчины",

xlab = "Места",

ylab = "Количество")

years\_woman <- df\_woman$Год[df\_woman[, 2] == 1]

years\_mix <- df\_mix$Год[df\_mix[, 2] == 1]

years\_man <- df\_man$Год[df\_man[, 2] == 1]

# Данные для первых мест

first\_place\_woman <- sum(df\_woman[,2] == 1)

first\_place\_mix <- sum(df\_mix[,2] == 1)

first\_place\_man <- sum(df\_man[,2] == 1)

# Круговая диаграмма

par(mfrow = c(1, 3))

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

if (first\_place\_woman == 0) {

pie(c(1, 0), labels = first\_place\_woman, main = "Женщины", col = "gray")

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

} else {

pie(c(first\_place\_woman, 0), labels = first\_place\_woman, main = "Женщины", col = legend\_colors)

legend("topright", legend = paste(years\_woman, collapse = ", "), fill = legend\_colors)

}

if (first\_place\_mix == 0) {

pie(c(1, 0), labels = first\_place\_mix, main = "Микс", col = "gray")

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

} else {

pie(c(first\_place\_mix, 0), labels = first\_place\_mix, main = "Микс", col = legend\_colors)

legend("topright", legend = paste(years\_mix, collapse = ", "), fill = legend\_colors)

}

if (first\_place\_man == 0) {

pie(c(1, 0), labels = first\_place\_man, main = "Мужчины", col = "gray")

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

} else {

pie(c(first\_place\_man, 0), labels = first\_place\_man, main = "Мужчины", col = legend\_colors)

legend("topright", legend = paste(years\_man, collapse = ", "), fill = legend\_colors)

}

# Функциональные графики

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]+0.5)))

axis(1, at = df$Год, labels = df$Год)

}

par(mfrow = c(3, 1))

plot\_graph(df\_woman, "red", "Тенденции изменения кол-ва призовых мест для женщин за последние 30 лет")

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

plot\_graph(df\_mix, "darkgreen", "Тенденции изменения кол-ва призовых мест для микса за последние 30 лет")

# Графики изменения спортивных достижений (золото)

df\_gold <- read\_excel("Olimp.xlsx", sheet = 4)

par(mar = c(5, 4, 4, 9))

matplot(df\_gold$Год,

df\_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,

col = legend\_colors,

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

xaxt = "n",

ylim = c(0, max(df\_gold[, -1]+10)))

axis(1, at = df\_gold$Год, labels = df$df\_gold)

legend("topright", legend = colnames(df\_gold)[-1], xpd = TRUE, bty = "n", inset = -c(0.35, 0), fill = legend\_colors)

# Графики изменения спортивных достижений (все призовые)

df\_prize\_places <- read\_excel("Olimp.xlsx", sheet = 5)

par(mar = c(5, 4, 4, 9))

matplot(df\_prize\_places$Год,

df\_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,

col = legend\_colors,

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

xaxt = "n",

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

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

legend("topright", legend = colnames(df\_prize\_places)[-1], xpd = TRUE, bty = "n", inset = -c(0.35, 0), fill = legend\_colors)

# Динамика фигурного катания по всем за последние 6 Олимпиад

all <- cbind(df\_mix[, 1:4], df\_woman[,2:4], df\_man[,2:4])

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

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

total\_places3 <- rowSums(df\_man[, 2:4], na.rm = TRUE)

total\_places\_all <- cbind(df\_mix[, 1], total\_places1, total\_places2 + 0.01, total\_places3)

year <- all[1:6, ]

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

matplot(total\_places\_all$Год,

total\_places\_all,

type = "b", pch = 19,

lty = c(1, 6, 3),

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

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

xaxt = "n",

lwd = 3,

ylim = c(0, max(year[, -1])+0.5))

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

legend('topleft', c('Мужчины', 'Микс', 'Женщины'), lty = c(1, 6, 3), cex = 0.8, lwd=4, col = c("blue", "darkgreen", "red"))

total\_places\_barplot <- cbind(df\_mix[, 1], total\_places1, total\_places2, total\_places3)

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

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

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

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

beside = TRUE,

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

xlab = "Год",

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

ylab = "Количество мест (призовые)",

ylim = c(0,3),

cex.names = 0.8)

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

total\_places\_pie <- sapply(total\_places\_barplot[,-1], sum)

pie(total\_places\_pie,

labels=c(total\_places\_pie[1], total\_places\_pie[2], total\_places\_pie[3]),

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

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

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

mtext("Общее количество призовых мест на 6 Олимпиадах по фигурному катанию",

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