install.packages("ggplot2", "dplyr", "GGally", "reshape2", "patchwork")

library(ggplot2)

library(dplyr)

library(GGally)

library(reshape2)

library(patchwork)

data <- read.csv('C:/Users/gelzi/Desktop/Duomenų mokslas/4 semestras/Duomenų vizualizavimas/2 lab/breast-cancer-wisconsin.csv') #nuskaitome duomenis

data

df1 <- data[c(2,6,8)]

colnames(df1) <- c('Auglys','Plotas','Įdubimas')

p1 <- ggplot(df1, aes(x=Įdubimas, y=Plotas, color = Auglys)) +

geom\_point(alpha=0.7) +

ggtitle('Ploto ir įdubimo sklaidos diagrama') +

scale\_color\_brewer(palette = "Set2", direction = 1)

p2 <- ggplot(data, aes(x = area, fill = Diagnosis)) +

geom\_histogram(position = "identity", alpha = 0.4, binwidth = 30, color="black") +

scale\_x\_continuous(breaks = seq(200, 2500, 200)) +

ggtitle("Piktybio ir nepiktybinio auglio ploto pasiskirstymas") +

labs(y = 'Dažnis', x = 'Plotas', fill = 'Auglys') +

scale\_fill\_brewer(palette = "Set2", direction = 1) +

scale\_y\_continuous(breaks = seq(0, 40, 5))

df <- data[c(2,3,4,5)] #stulpeliu pavadinimai

colnames(df) <- c('Auglys','spindulys', 'tekstūra', 'perimetras')

# scale - globalminmax -> parodo tikrąsias reikšmes grafike

# alphaLines -> nustato linijų permatomumą

# groupColumn -> pagal ką grupuosime duomenis

# M - piktybinis auglys (malignant)

# B - neiktybinis auglys (benign)

p3 <- ggparcoord(data = df, columns = 2:4, groupColumn = 1, scale = "globalminmax",

showPoints = TRUE , alphaLines = 0.4, title = "Lygiagrečiųjų koordinačių grafikas apie krūtų vėžį") +

theme(plot.title = element\_text(size=13)) +

xlab(' ') + ylab('Skaičius') +

scale\_color\_brewer(palette = "Set2")

p <- (p1 / p3) | p2 + plot\_annotation(tag\_levels = "I")

p + plot\_annotation(title = "Piktybinio ir nepiktybinio krūtų vėžio auglio charakteristikų palyginimas", theme = theme(plot.title = element\_text(hjust = 0.5, size = 18)))