



**Vilniaus  
universitetas**

# MATEMATIKOS IR INFORMATIKOS FAKULTETAS

Simona Gelžinytė

## **5 laboratorinis darbas**

Darbo aprašas

Vilnius, 2022

# Turiny

<b>Įvadas .....</b>	<b>3</b>
<b>Tikslas .....</b>	<b>3</b>
<b>Uždaviniai .....</b>	<b>3</b>
<b>Tiriamajo darbo pagrindinė dalis.....</b>	<b>3</b>
<b>Pasirinkti duomenys .....</b>	<b>3</b>
<b>Pirmas grafikas .....</b>	<b>3</b>
<b>Antras grafikas.....</b>	<b>5</b>
<b>Šaltiniai .....</b>	<b>6</b>

## Ivadas

Tai yra duomenų vizualizavimo laboratorinio darbo aprašas, rašytas Vilniaus Universiteto studentės Simonos Gelžinytės. Šiame darbe pateikti 2 interaktyvūs grafikai, naudojant „R“ programos paketus – „ggplot2“, „plotly“. Užduotis atlikta naudojant duomenis iš „UCI Machine Learning Repository“ apie krūtų vėžį.

## Tikslas

Turint duomenų rinkinį apie krūties vėžio atvejus Viskonsino valstijoje, JAV palyginti invazinių ir neinvazinių krūtų vėžio auglio charakteristikas.

## Uždaviniai

- Išsiaiškinti koks ryšys yra tarp auglio perimetro ir tekstūros
- Išsiaiškinti koks ryšys yra tarp auglio perimetro ir spindulio.

## Tiriamajo darbo pagrindinė dalis

### Pasirinkti duomenys

Duomenis parinkti iš „UCI Machine Learning Repository“ apie krūtų vėžį Viskonsino valstijoje. Vertinta pagal auglio perimetrą, spindulį ir tekstūrą.

### Pirmas grafikas

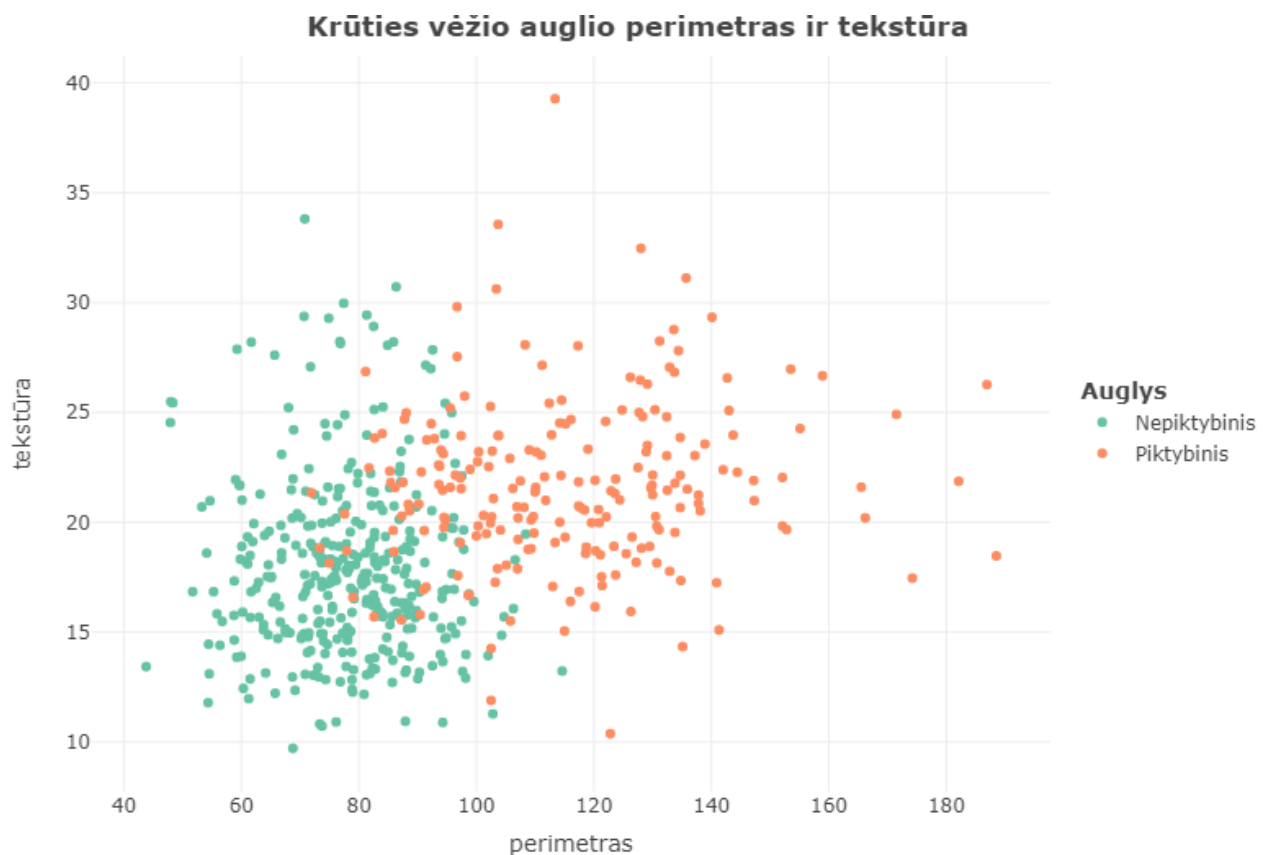
```
#stulpeliai, kuriuos naudosime
df <- data[c(2,3,4,5)]
#stulpelių pavadinimai
colnames(df) <- c('Auglys', 'spindulys', 'tekstūra', 'perimetras')
#data - pasirinkti duomenys, x - x asyje pasirinktas stulpelis, y - y asyje
pasirinktas stulpelis
#color - pagal kuri stulpeli parinkti spalvas (t.y. kiek yra skirtingu
kintamųjų, kad zinoti kiek bus skirtingu spalvu)
```

```
#colors - spalvu rinkinys

fig <- plot_ly(data = df, x = ~perimetras, y = ~tekstūra, color = ~Auglys,
  colors = "Set2")

#title - grafiko pavadinimas, legend title - legendos pavadinimas, legend list -
nustatome koordinates, kad legena butu viduryje

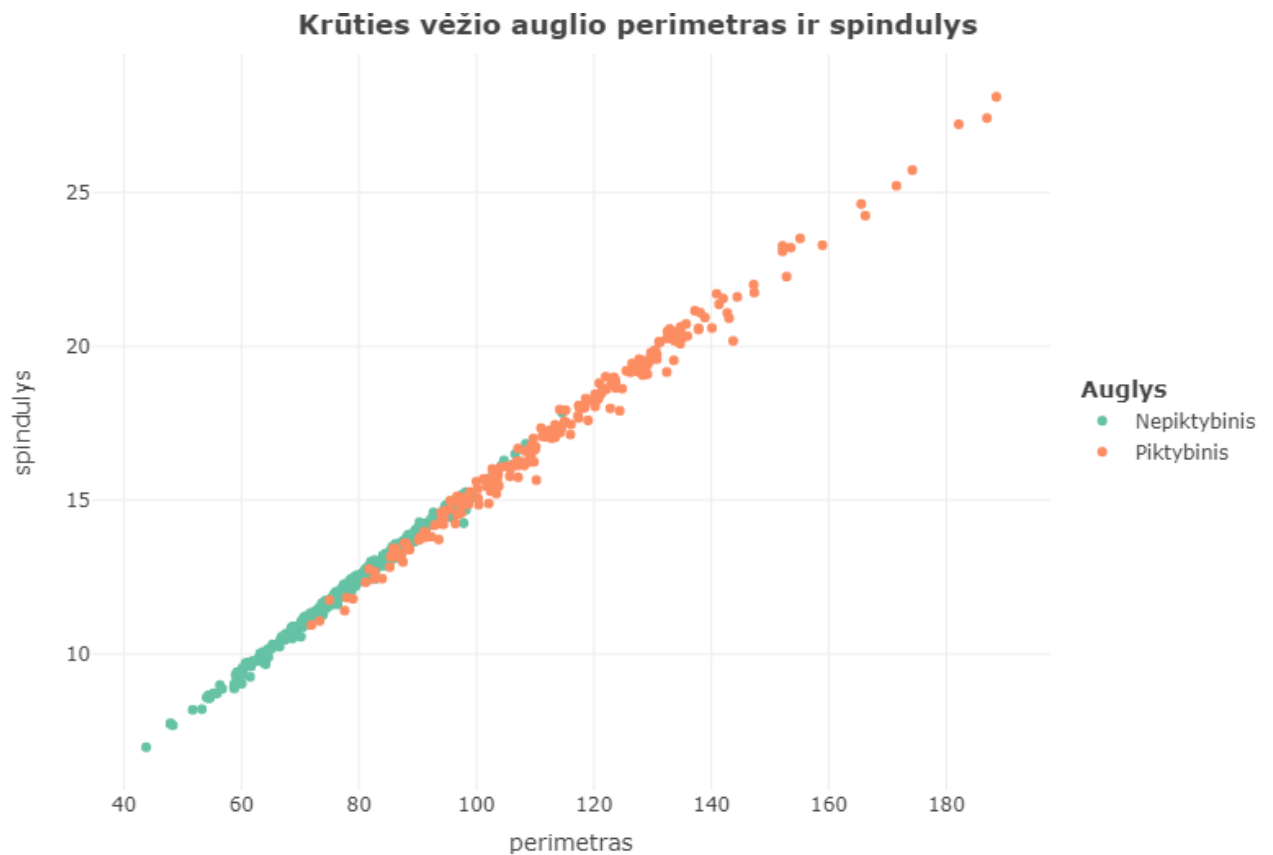
fig <- fig %>% layout(title = '<b> Krūties vėžio auglio perimetras ir tekstūra
<b>', legend=list(title=list(text='<b> Auglys </b>'), x=100, y =0.5))
```



**Išvada:** iš grafiko matome, jog invazinių auglių perimetras yra pasiskirstęs labiau dešinėje pusėje, tai reiškia didesnę perimetrą, bet tiesioginio ryšio tarp perimetro ir tekstūros negalime matyti (t.y. nebūtinai kuo didesnis perimetras, tuo kietesnis auglys yra).

## Antras grafikas

```
fig2 <- plot_ly(data = df, x = ~perimetras, y = ~spindulys, color = ~Auglys,  
colors = "Set2")  
fig2 <- fig2 %>% layout(title = '<b> Krūties vėžio auglio perimetras ir  
spindulys <b>', legend=list(title=list(text='<b> Auglys </b>'), x=100, y =0.5))
```



**Išvada:** iš grafiko matome, jog piktybinių auglių perimetras ir spindulys yra didesni už nepiktybinių. Taip pat galime pasakyti, jog šios charakteristikos turi tiesioginį ryšį (t.y. kuo didesnis spindulys, tuo didesnis ir perimetras).

## Šaltiniai

### Naudota literatūra

- <https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29>
- <https://plotly.com/r/basic-charts/>