Błażej Domagała - WED

# Lab 1

## Zadanie 1

a) Nazwa zbioru: Wine

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

b) **Krótki tekstowy opis zbioru**

Zbiór danych "Wine" zawiera wyniki analizy chemicznej win wyprodukowanych w określonym regionie we Włoszech przez trzech różnych producentów. Analiza chemiczna dotyczy 13 różnych składników zawartych w winach.

Obraz zawierający zrzut ekranu, tekst, Czcionka

Opis wygenerowany automatycznie

c) Liczba obserwacji w zbiorze: 178

Obraz zawierający tekst, Czcionka, zrzut ekranu, numer

Opis wygenerowany automatycznie

d) Liczba kolumn: 14 (13 atrybutów + 1 kolumna identyfikująca klasę)

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

e) Zmienna celu:

* Nazwa kolumny z klasą: Class
* Liczba klas: 3 (Klasy 1, 2 i 3)

Obraz zawierający tekst, Czcionka, zrzut ekranu, numer

Opis wygenerowany automatycznie

f) Wykaz i opis cech:

1. **Class**: Zmienna kategoryczna. Klasa wina.
2. **Alcohol**: Zmienna ilościowa. Zawartość alkoholu.
3. **Malic acid**: Zmienna ilościowa. Zawartość kwasu jabłkowego.
4. **Ash**: Zmienna ilościowa. Zawartość popiołu.
5. **Alcalinity of ash**: Zmienna ilościowa. Zasadowość popiołu.
6. **Magnesium**: Zmienna ilościowa. Zawartość magnezu.
7. **Total phenols**: Zmienna ilościowa. Całkowita zawartość fenoli.
8. **Flavanoids**: Zmienna ilościowa. Zawartość flawonoidów.
9. **Nonflavanoid phenols**: Zmienna ilościowa. Zawartość fenoli nieflawonoidowych.
10. **Proanthocyanins**: Zmienna ilościowa. Zawartość proantocyjanidyn.
11. **Color intensity**: Zmienna ilościowa. Intensywność koloru.
12. **Hue**: Zmienna ilościowa. Odcień.
13. **OD280/OD315 of diluted wines**: Zmienna ilościowa. Stosunek absorbancji przy 280 nm do 315 nm w rozcieńczonych winach.
14. **Proline**: Zmienna ilościowa. Zawartość prolina.

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

## Zadanie 2

1. Zmiana nazw kolumn (pierwszą kolumnę – zmienną celu – proszę nazwać „Class”; nazwy pozostałych kolumn – atrybutów – są podane w pliku „wine.names”)

Obraz zawierający tekst, oprogramowanie, Oprogramowanie multimedialne, zrzut ekranu

Opis wygenerowany automatycznie

1. Polecenie View (fragment print-screena z tabelką)

Obraz zawierający Czcionka, tekst, zrzut ekranu, linia

Opis wygenerowany automatycznie

Obraz zawierający zrzut ekranu, tekst, wyświetlacz, oprogramowanie

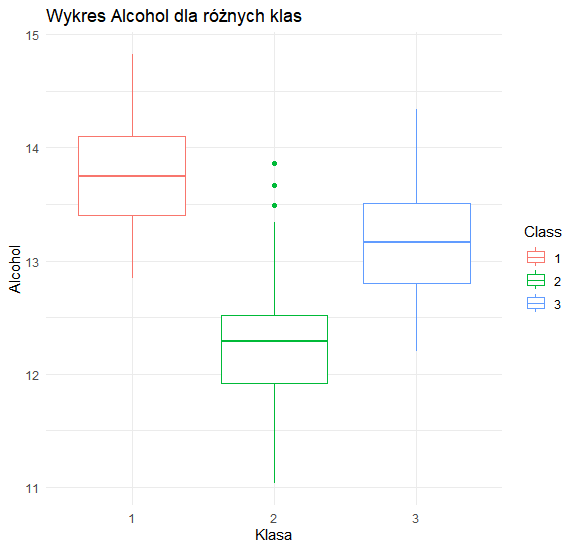
Opis wygenerowany automatycznie

1. Podsumowanie cech (summary)

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

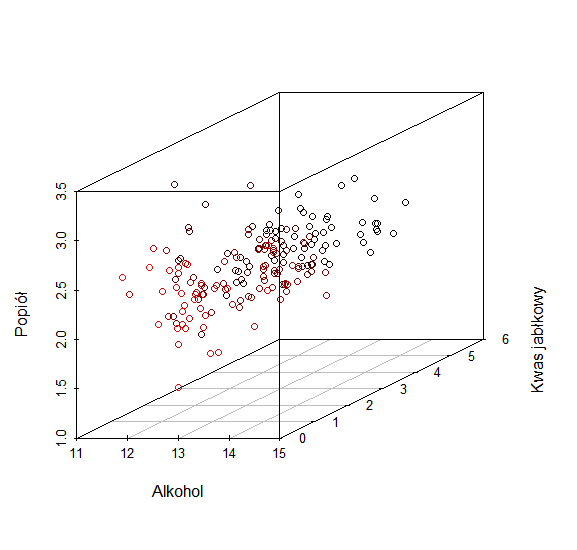
1. Wykres 2D ilustrujący wybraną cechę dla różnych klas



Obraz zawierający tekst, zrzut ekranu

Opis wygenerowany automatycznie

1. Wykres 3D dla trzech wybranych cech (bez klasy)



Obraz zawierający tekst, Czcionka, zrzut ekranu

Opis wygenerowany automatycznie

# Lab 2

1. Dla wybranej cechy wyświetlić wybrane wartości stosując: zakres wartości, sekwencję indeksów (np. co dziesiąty indeks), indeksy ujemne, warunki logiczne.

Kod:

path = "C:\\Users\\petitoff\\Desktop\\repos\\UO\\rok 3\\Wprowadzenie do eksploracji danych\\lista2"

setwd(path) ## ustawienie ścieżki

# Załadowanie danych

wine <- read.csv('wine\\wine.data', header = FALSE)

# Zmień nazwy kolumn

names(wine) <-

c(

'Class',

'Alcohol',

'Malic acid',

'Ash',

'Alcalinity of ash',

'Magnesium',

'Total phenols',

'Flavanoids',

'Nonflavanoid phenols',

'Proanthocyanins',

'Color intensity',

'Hue',

'OD280/OD315 of diluted wines',

'Proline'

)

# a) Dla wybranej cechy wyświetlić wybrane wartości stosując: zakres wartości, sekwencję indeksów (np. co dziesiąty indeks), indeksy ujemne, warunki logiczne.

# Zakres wartości

wine$Alcohol[5:15]

# Sekwencja indeksów

wine$Alcohol[seq(1, nrow(wine), 10)]

# Indeksy ujemne

wine$Alcohol[-(1:5)]

# Warunki logiczne

wine$Alcohol[wine$Alcohol > 14]

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

1. Wyświetlić wybrane wiersze i kolumny z tabeli.

# b) Wybranie wierszy i kolumn:

selected\_rows\_columns <- wine[1:10, c("Alcohol", "Malic acid")]

print(selected\_rows\_columns)

# indeksy

wine[5:15,]

Obraz zawierający tekst, zrzut ekranu

Opis wygenerowany automatycznie

1. Dodać do tabeli nową kolumnę z wartościami obliczonymi na podstawie innych wybranych kolumn.

# c) Dodanie nowej kolumny:

# Sprawdzanie, czy kolumna istnieje i zawiera dane

if ("Total phenols" %in% names(wine) &&

!all(is.na(wine$`Total phenols`))) {

# Dodanie nowej kolumny

wine$Total.phenols.squared <- wine$`Total phenols` ^ 2

head(wine)

} else {

cat("Column 'Total phenols' does not exist or is empty.")

}

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

1. Podać wartości podstawowych statystyk dla wybranej kolumny: zakres, średnia, mediana,

# d) Statystyki podstawowe dla wybranej kolumny, np. "Alcohol"

cat("Statystyki dla kolumny 'Alcohol':", "\n")

cat("Zakres: ", min(wine$Alcohol), " - ", max(wine$Alcohol), "\n")

cat("Średnia: ", mean(wine$Alcohol), "\n")

cat("Mediana: ", median(wine$Alcohol), "\n")

cat("Odchylenie standardowe: ", sd(wine$Alcohol), "\n")

cat("Kurtoza: ", moments::kurtosis(wine$Alcohol), "\n")

cat("Skośność: ", moments::skewness(wine$Alcohol), "\n")

cat("Kwantyle: ", quantile(wine$Alcohol, probs = c(0.25, 0.5, 0.75)), "\n")

Obraz zawierający tekst, zrzut ekranu, Czcionka, oprogramowanie

Opis wygenerowany automatycznie

1. Wyznaczyć i zilustrować na wykresie macierz korelacji dla wybranych pięciu zmiennych.

# e) Macierz korelacji dla wybranych pięciu zmiennych i jej wizualizacja

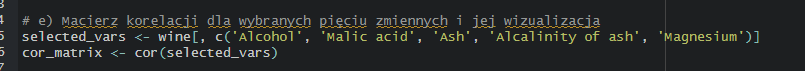
selected\_vars <-

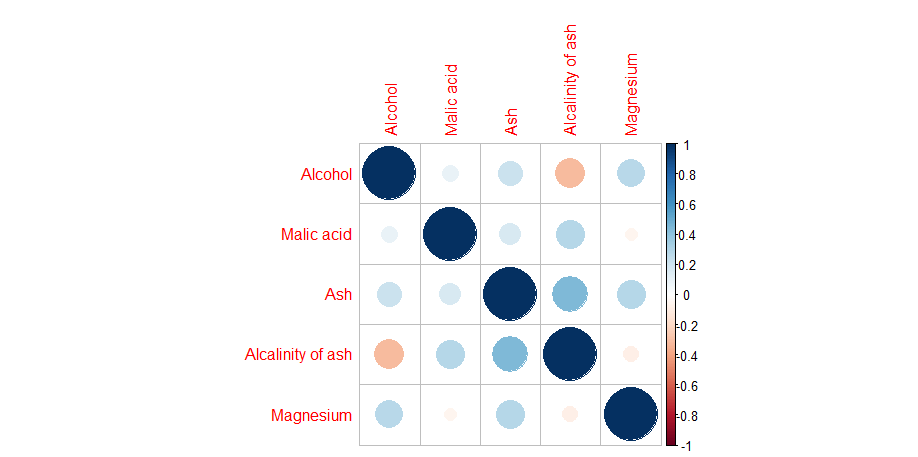
wine[, c('Alcohol', 'Malic acid', 'Ash', 'Alcalinity of ash', 'Magnesium')]

cor\_matrix <- cor(selected\_vars)

library(corrplot)

corrplot(cor\_matrix, method = "circle")





1. Wydrukować histogramy dla trzech różnych zmiennych, przedyskutować wyniki.

# f) Histogramy dla trzech różnych zmiennych

par(mfrow = c(1, 3)) # ustawienie layoutu na 1 wiersz i 3 kolumny

hist(wine$Alcohol,

main = 'Alcohol',

xlab = '',

col = 'skyblue')

hist(

wine$`Malic acid`,

main = 'Malic acid',

xlab = '',

col = 'skyblue'

)

hist(wine$Ash,

main = 'Ash',

xlab = '',

col = 'skyblue')

Obraz zawierający tekst, zrzut ekranu, Czcionka, oprogramowanie

Opis wygenerowany automatycznie

