

Федеральное государственное бюджетное образовательное
учреждение
высшего образования
«Саратовский государственный технический университет
имени Гагарина Ю.А.»

Кафедра «Прикладные информационные технологии»

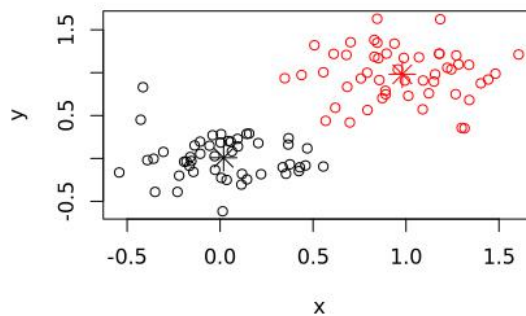
ОТЧЕТ
по лабораторной работе номер 7

Студента гр. 62-ПИНФ21:
Нефедова Данила Вадимовича
Проверил доцент кафедры ПИТ:
Бровко Александр Валерьевич

Саратов 2019

1.2

```
#  
# 1. 2. Воспроизведите пример, приведенный в справке, для случая сгенерированных  
#       данных с двумя кластерами.  
#  
x <- rbind(matrix(rnorm(100, sd = 0.3), ncol = 2), matrix(rnorm(100, mean = 1, sd = 0.3),  
ncol = 2))  
colnames(x) <- c("x", "y")  
cl <- kmeans(x, 2)  
plot(x, col = cl$cluster)  
points(cl$centers, col = 1:2, pch = 8, cex = 2)
```



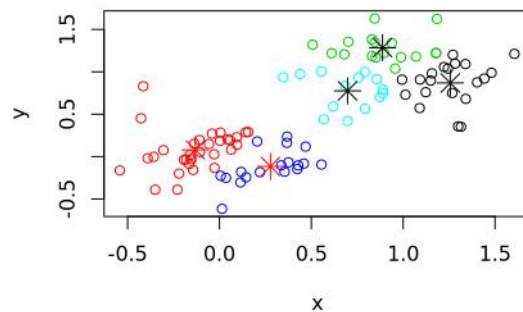
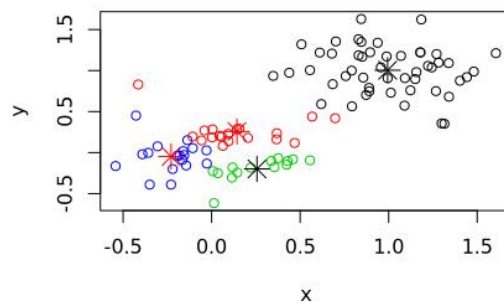
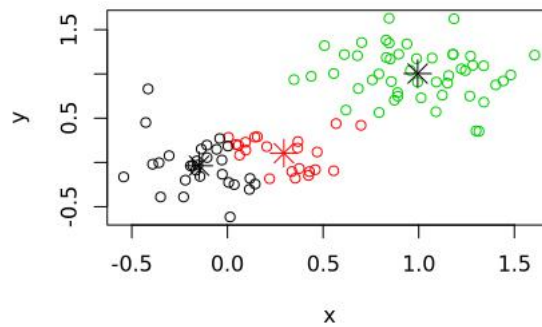
1.3

```
#  
# 1. 3. Выполните кластеризацию для сгенерированных в пункте 1.2 данных,  
#       для случаев K = 3, 4, 5. Изобразите кластеры с их центрами на  
#       отдельных рисунках, сохраните их в графическом формате,  
#       вставьте в отчет.  
#
```

```
cl <- kmeans(x, 3)  
plot(x, col = cl$cluster)  
points(cl$centers, col = 1:2, pch = 8, cex = 2)
```

```
cl <- kmeans(x, 4)  
plot(x, col = cl$cluster)  
points(cl$centers, col = 1:2, pch = 8, cex = 2)
```

```
cl <- kmeans(x, 5)  
plot(x, col = cl$cluster)  
points(cl$centers, col = 1:2, pch = 8, cex = 2)
```



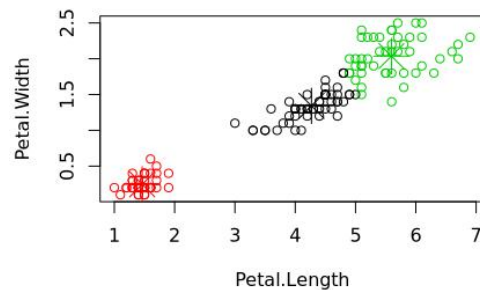
1.4

```
#
# 1. 4. Посмотрите структуру набора данных iris, поставляемого вместе с R.
#      Выполните кластеризацию K-средних для набора iris, используя в качестве
#      данных значения Petal.Length и Petal.Width (3 и 4 столбцы набора данных
#      iris). В качестве K возьмите значение 3.
#
```

```
x <- iris[, 3:4]
```

```
cl <- kmeans(x, 3)
plot(x, col = cl$cluster)
points(cl$centers, col = 1:3, pch = 8, cex = 3)
```

1.5



1.6

```
#
# 1.6. Сделайте график зависимости WSS от числа кластеров,
#       используя следующий код:
#
```

```
dat <- iris[, 3:4]
```

```
mydata <- dat
```

```
wss <- (nrow(mydata) - 1) * sum(apply(mydata, 2, var))
```

```
for (i in 2:15) wss[i] <- sum(kmeans(mydata, centers=i)$withinss)
```

```
plot(1:15,
     wss,
     type = "b",
     xlab = "Number of Clusters",
     ylab = "Within groups sum of squares",
     main = "Assessing the Optimal Number of Clusters with the Elbow Method",
     pch = 20,
     cex = 2)
```

