

# Testy

Michał Milewski

5/14/2020

## Dane użyte do testów

Do testów zostały spreparowane 3 zestawy danych. Wszystkie one składają się z 3 grup punktów, oetyki-  
etowanych 3 różnymi etykietami. W każdej z grup znajduje się po 100 punktów. Główną różnicą między  
tymi zbiorami jest zgęszczenie punktów w grupach i częściowe zachodzenie na siebie obszarów różnych grup.

```
num_of_points <- 100
```

- pierwszy zbiór

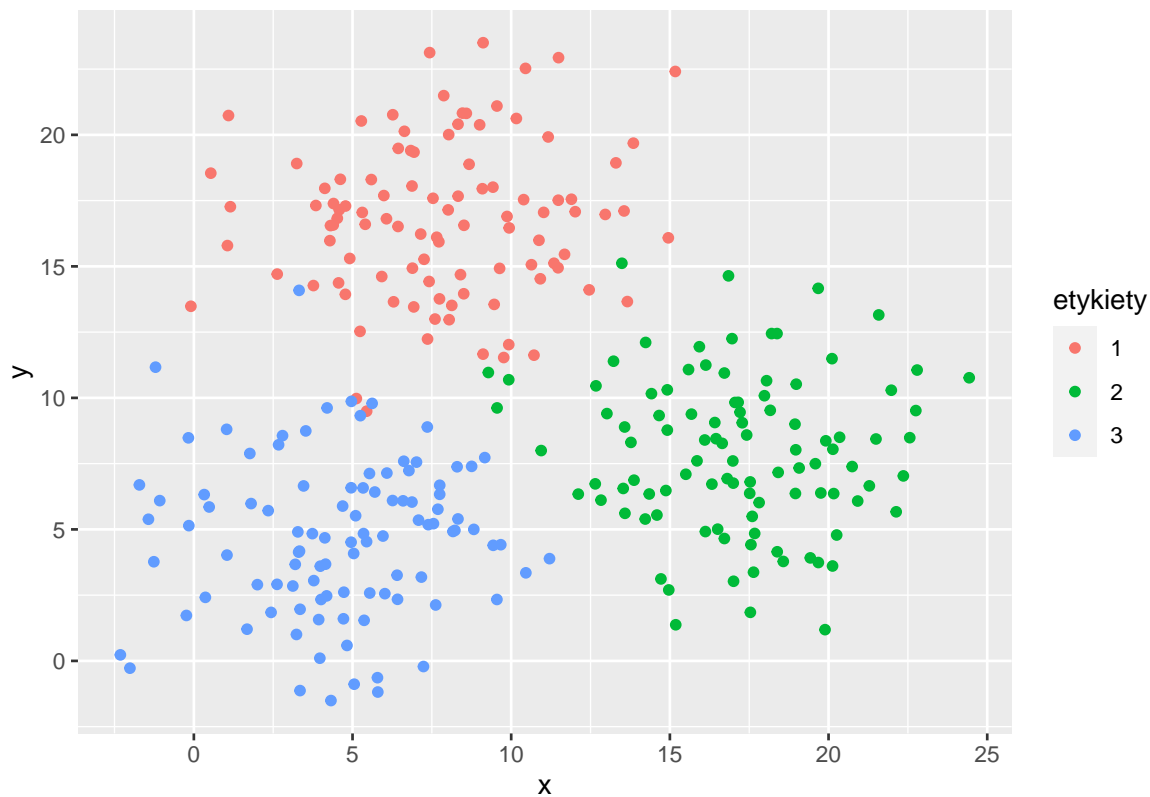
```
first_set <- cbind(rnorm(num_of_points,mean = 8,sd = 3),rnorm(num_of_points,mean = 17,sd = 3))
first_set_etykiety <- rep(1,num_of_points)

tmp_matrix <- cbind(rnorm(num_of_points,mean = 17,sd = 3),rnorm(num_of_points,mean = 8,sd = 3))
tmp_etykiety <- rep(2,num_of_points)

first_set <- rbind(first_set,tmp_matrix)
first_set_etykiety <- c(first_set_etykiety,tmp_etykiety)

tmp_matrix <- cbind(rnorm(num_of_points,mean = 5,sd = 3),rnorm(num_of_points,mean = 5,sd = 3))
tmp_etykiety <- rep(3,num_of_points)

first_set <- rbind(first_set,tmp_matrix)
first_set_etykiety <- c(first_set_etykiety,tmp_etykiety)
```



- drugi zbiór

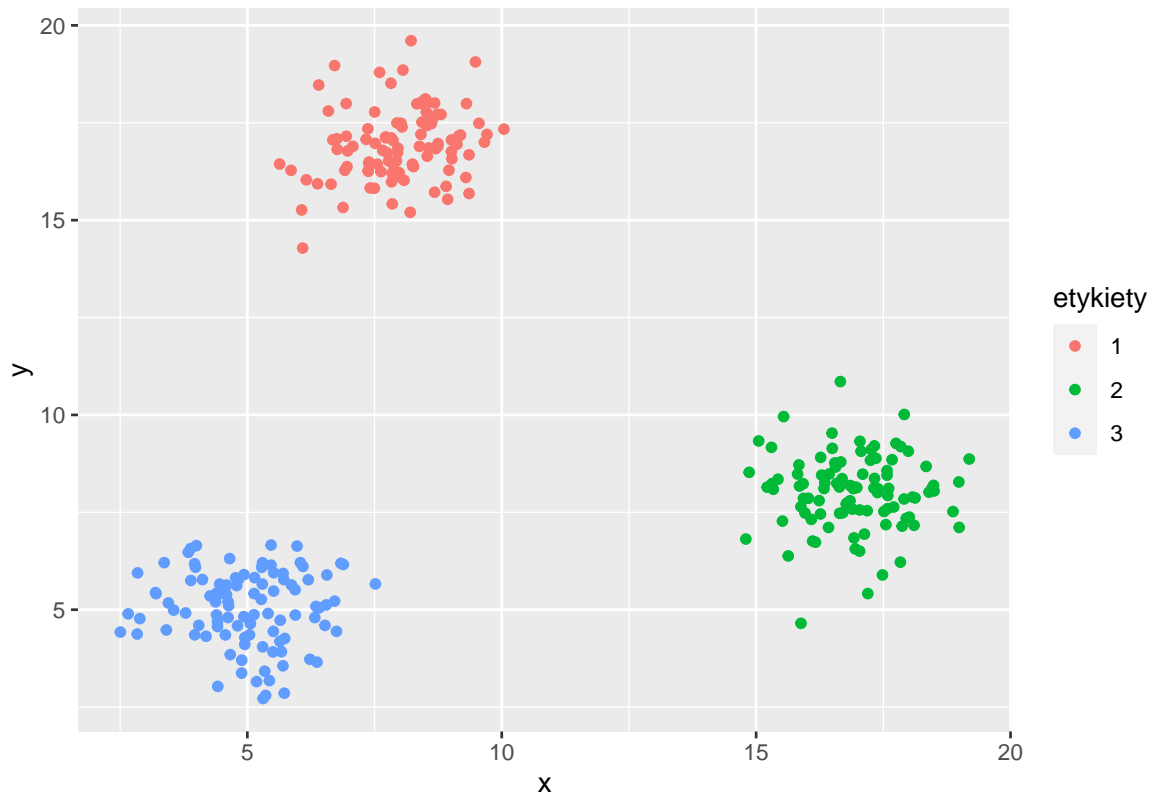
```
second_set <- cbind(rnorm(num_of_points,mean = 8,sd = 1),rnorm(num_of_points,mean = 17,sd = 1))
second_set_etykiety <- rep(1,num_of_points)

tmp_matrix <- cbind(rnorm(num_of_points,mean = 17,sd = 1),rnorm(num_of_points,mean = 8,sd = 1))
tmp_etykiety <- rep(2,num_of_points)

second_set <- rbind(second_set,tmp_matrix)
second_set_etykiety <- c(second_set_etykiety,tmp_etykiety)

tmp_matrix <- cbind(rnorm(num_of_points,mean = 5,sd = 1),rnorm(num_of_points,mean = 5,sd = 1))
tmp_etykiety <- rep(3,num_of_points)

second_set <- rbind(second_set,tmp_matrix)
second_set_etykiety <- c(second_set_etykiety,tmp_etykiety)
```



- trzeci zbiór

```
third_set <- cbind(rnorm(num_of_points,mean = 8,sd = 1),rnorm(num_of_points,mean = 17,sd = 1))
third_set_etykiety <- rep(1,num_of_points)

tmp_matrix <- cbind(rnorm(num_of_points,mean = 17,sd = 3),rnorm(num_of_points,mean = 8,sd = 3))
tmp_etykiety <- rep(2,num_of_points)

third_set <- rbind(third_set,tmp_matrix)
third_set_etykiety <- c(third_set_etykiety,tmp_etykiety)

tmp_matrix <- cbind(rnorm(num_of_points,mean = 5,sd = 5),rnorm(num_of_points,mean = 5,sd = 5))
tmp_etykiety <- rep(3,num_of_points)

third_set <- rbind(third_set,tmp_matrix)
third_set_etykiety <- c(third_set_etykiety,tmp_etykiety)
```



- punkty testowe, dla których będą znajdowane etykiety

x	y	etykiety
5	5	3
17	8	2
8	17	1
0	0	3
8	22	1
22	8	2
5	5	3
17	8	2
8	17	1
0	0	3
8	22	1
22	8	2

## Podstawowe testy poprawności działania 1-nn

Podstawowy test polega na podaniu zarówno jako zbioru uczącego jak i zbioru testowego. Następnie następuje porównanie czy etykiety wybrane są równe etykietom ze zbioru uczącego.

```
res <- knn(first_set,first_set_etykiety,first_set,1,1)
all.equal(res,first_set_etykiety)
```

```
## [1] TRUE
```

```

res <- knn(first_set,first_set_etykiety,first_set,1,2)
all.equal(res,first_set_etykiety)

## [1] TRUE

res <- knn(first_set,first_set_etykiety,first_set,1,Inf)
all.equal(res,first_set_etykiety)

## [1] TRUE

res <- knn(second_set,second_set_etykiety,second_set,1,1)
all.equal(res,second_set_etykiety)

## [1] TRUE

res <- knn(second_set,second_set_etykiety,second_set,1,2)
all.equal(res,second_set_etykiety)

## [1] TRUE

res <- knn(second_set,second_set_etykiety,second_set,1,Inf)
all.equal(res,second_set_etykiety)

## [1] TRUE

res <- knn(third_set,third_set_etykiety,third_set,1,1)
all.equal(res,third_set_etykiety)

## [1] TRUE

res <- knn(third_set,third_set_etykiety,third_set,1,2)
all.equal(res,third_set_etykiety)

## [1] TRUE

res <- knn(third_set,third_set_etykiety,third_set,1,Inf)
all.equal(res,third_set_etykiety)

## [1] TRUE

```

## Podstawowe testy działania funkcji agregujących

### Moda

Prosty test sprawdzający poprawne wybranie mody z każdego wiersza macierzy

```

testing <- matrix(c(1,2,2,1,2,
                    3,1,1,3,3,
                    6,7,3,4,7,
                    9,0,9,8,1),ncol = 5, byrow = TRUE)
res <- moda(testing)
all.equal(res,c(2,3,7,9))

## [1] TRUE

```

### Średnia

Prosty test sprawdzający liczenie średniej z każdego wiersza

```
testing <- matrix(c(1,2,2,1,2,
                   3,1,1,3,3,
                   6,7,3,4,7,
                   9,0,9,8,1),ncol = 5, byrow = TRUE)
res <- moda(testing)
all.equal(res,c(2,2,5,5))

## [1] "Mean relative difference: 0.3684211"
```

### Minkara1.5

Prosty test sprawdzający liczenie minkara1.5 z każdego wiersza

```
testing <- matrix(c(1,2,2,1,2,
                   3,1,1,3,3,
                   6,7,3,4,7,
                   9,0,9,8,1),ncol = 5, byrow = TRUE)
res <- minkara1.5(testing)
all.equal(res,c(2,2,6,6))

## [1] TRUE
```

### Minkara1.5

Prosty test sprawdzający liczenie minkara3.0 z każdego wiersza

```
testing <- matrix(c(1,2,2,1,2,
                   3,1,1,3,3,
                   6,7,3,4,7,
                   9,0,9,8,1),ncol = 5, byrow = TRUE)
res <- minkara3.0(testing)
all.equal(res,c(2,2,5,5))

## [1] TRUE
```

### Średnia warzona

Prosty test sprawdzający liczenie średniej warzonej z każdego wiersza

```
testing <- matrix(c(1,2,2,1,2,
                   3,1,1,3,3,
                   6,7,3,4,7,
                   9,0,9,8,1),ncol = 5, byrow = TRUE)
res <- srednia_wazona(testing)
all.equal(res,c(2,2,5,6))

## [1] TRUE
```

## Testy przyporządkowywania etykiet

- Funkcja testująca
  - Funkcja ta używa knn dla podanych danych, następnie za pomocą zadanej funkcji agregującej

```

test_fun<- function(X, y, Z, k, p, aggregate_fun, poprawne_etykiety, draw_plot = FALSE){
  ERR <- 0
  MAD <- 0
  MSE <- 0
  X <- as.matrix(X)
  y <- as.vector(y)
  Z <- as.matrix(Z)
  res <- knn(X,y,Z,k,p)
  res <- aggregate_fun(res)
  ERR <- sum(ifelse((res - poprawne_etykiety)!=0,1,0))/length(res)
  MAD <- sum(abs(res - poprawne_etykiety))/length(res)
  MSE <- sum(abs(res - poprawne_etykiety)^2)/length(res)

  print("MAD:")
  print(MAD)
  print("MSE:")
  print(MSE)
  print("ERR:")
  print(ERR)

  set <- data.frame(x=X[,1],y=X[,2],etykieta=as.character(y))
  results <- data.frame(x=Z[,1],y=Z[,2],etykieta=as.character(res))
  pl <- qplot(x,y,data=set,colour = etykieta) + geom_point(data=results,shape = 18,size = 6)
  if(draw_plot){
    pl
  }
}

```

Pierwszy zbiór testowy

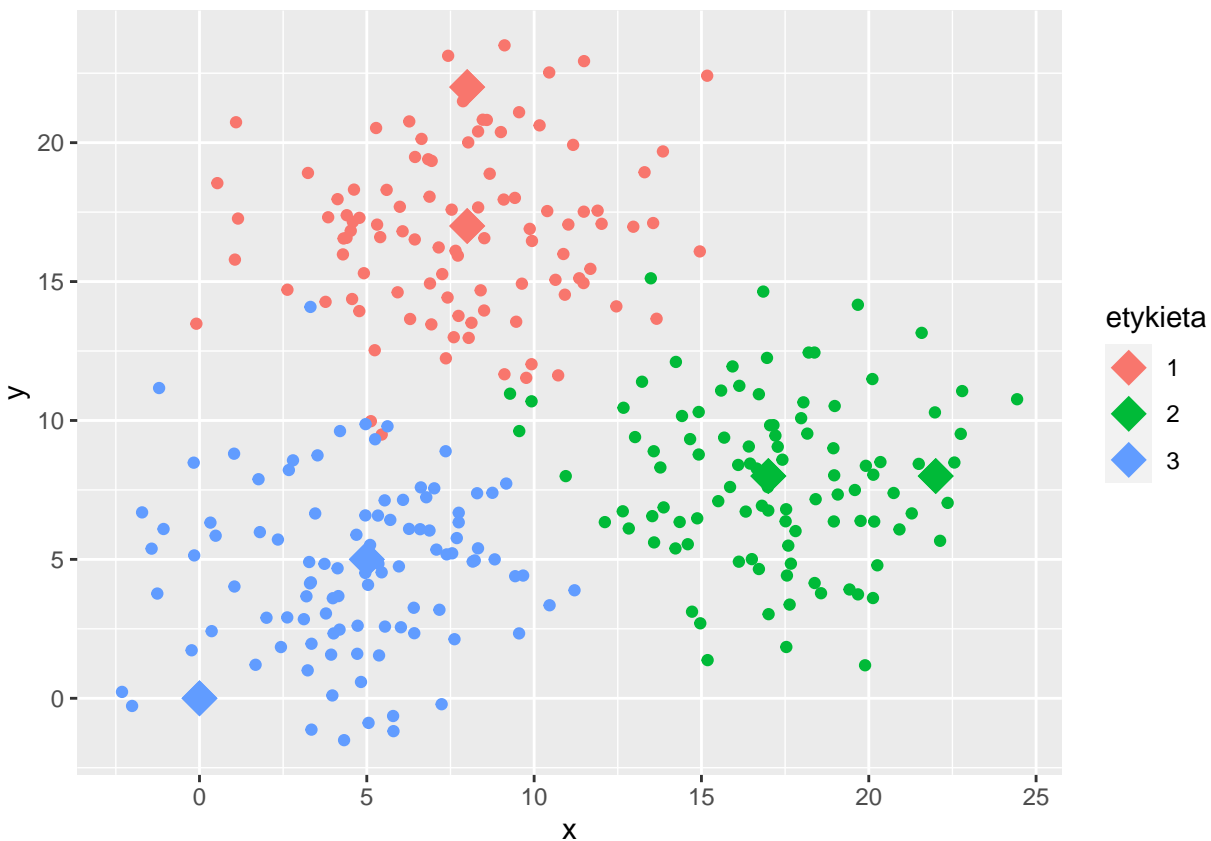
Funkcja agregująca moda

```

test_fun(first_set,first_set_etykiety,tested_points,19,2,moda,tested_points_etykiety,TRUE)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

```



```
test_fun(first_set,first_set_etykiety,tested_points,19,1,moda,tested_points_etykiety)
```

```
## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```

```
test_fun(first_set,first_set_etykiety,tested_points,19,Inf,moda,tested_points_etykiety)
```

```
## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```

Funkcja agregująca srednia\_a

```
test_fun(first_set,first_set_etykiety,tested_points,19,2,srednia_a,tested_points_etykiety)
```

```
## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
```



```
## [1] 0
test_fun(first_set,first_set_etykiety,tested_points,19,1,srednia_a,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

test_fun(first_set,first_set_etykiety,tested_points,19,Inf,srednia_a,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```

#### Funkcja agregująca minkara1.5

```
test_fun(first_set,first_set_etykiety,tested_points,19,2,minkara1.5,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

test_fun(first_set,first_set_etykiety,tested_points,19,1,minkara1.5,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

test_fun(first_set,first_set_etykiety,tested_points,19,Inf,minkara1.5,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```

#### Funkcja agregująca minkara3.0

```
test_fun(first_set,first_set_etykiety,tested_points,19,2,minkara3.0,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
```

```
## [1] "ERR:"
## [1] 0

test_fun(first_set,first_set_etykiety,tested_points,19,1,minkara3.0,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

test_fun(first_set,first_set_etykiety,tested_points,19,Inf,minkara3.0,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```

#### Funkcja agregująca srednia\_wazona

```
test_fun(first_set,first_set_etykiety,tested_points,19,2,srednia_wazona,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

test_fun(first_set,first_set_etykiety,tested_points,19,1,srednia_wazona,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

test_fun(first_set,first_set_etykiety,tested_points,19,Inf,srednia_wazona,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```

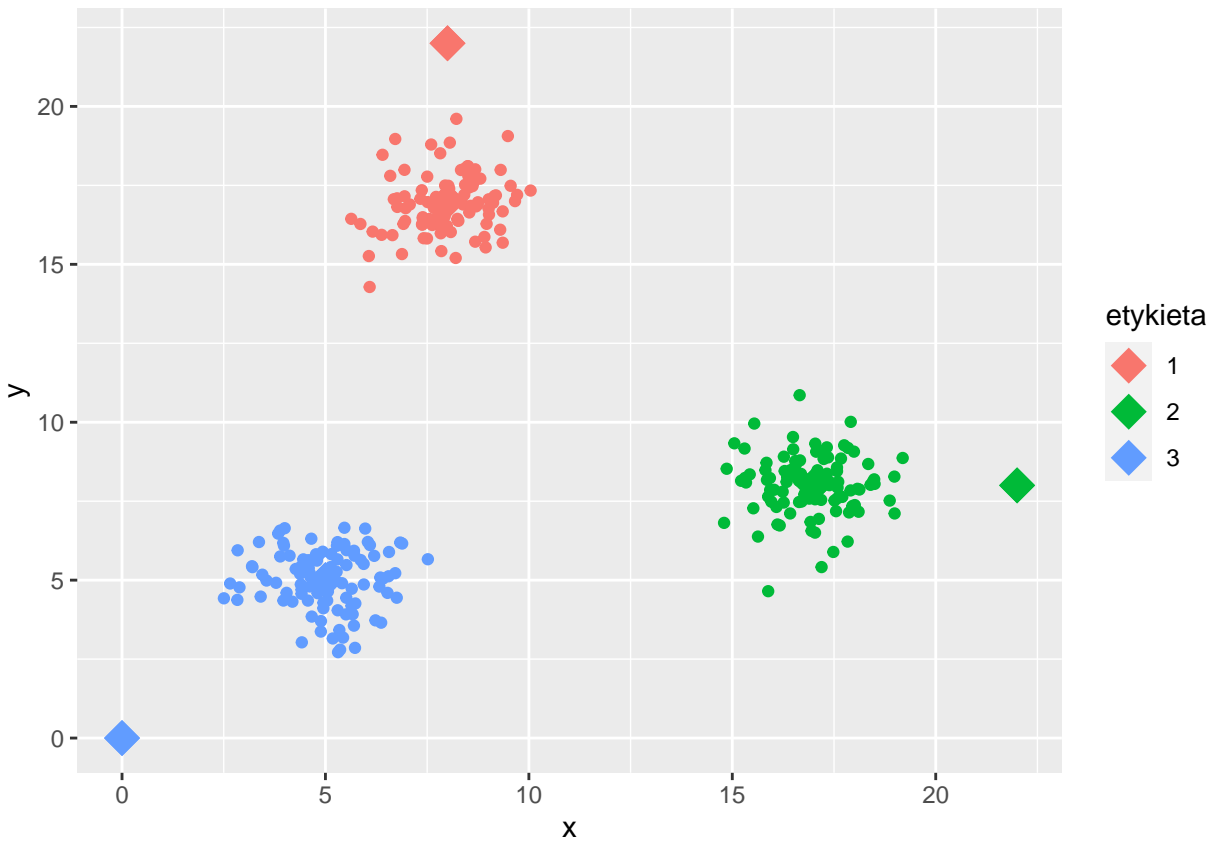
#### Drugi zbiór testowy

##### Funkcja agregująca moda

```
test_fun(second_set,second_set_etykiety,tested_points,19,2,moda,tested_points_etykiety,TRUE)

## [1] "MAD:"
```

```
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```



```
test_fun(second_set,second_set_etykiety,tested_points,19,1,moda,tested_points_etykiety)
```

```
## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```

```
test_fun(second_set,second_set_etykiety,tested_points,19,Inf,moda,tested_points_etykiety)
```

```
## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```

Funkcja agregująca srednia\_a

```

test_fun(second_set,second_set_etykiety,tested_points,19,2,srednia_a,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

test_fun(second_set,second_set_etykiety,tested_points,19,1,srednia_a,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

test_fun(second_set,second_set_etykiety,tested_points,19,Inf,srednia_a,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

```

#### Funkcja agregująca minkara1.5

```

test_fun(second_set,second_set_etykiety,tested_points,19,2,minkara1.5,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

test_fun(second_set,second_set_etykiety,tested_points,19,1,minkara1.5,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

test_fun(second_set,second_set_etykiety,tested_points,19,Inf,minkara1.5,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

```

### Funkcja agregująca minkara3.0

```
test_fun(second_set,second_set_etykiety,tested_points,19,2,minkara3.0,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

test_fun(second_set,second_set_etykiety,tested_points,19,1,minkara3.0,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

test_fun(second_set,second_set_etykiety,tested_points,19,Inf,minkara3.0,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```

### Funkcja agregująca srednia\_wazona

```
test_fun(second_set,second_set_etykiety,tested_points,19,2,srednia_wazona,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

test_fun(second_set,second_set_etykiety,tested_points,19,1,srednia_wazona,tested_points_etykiety)

## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0

test_fun(second_set,second_set_etykiety,tested_points,19,Inf,srednia_wazona,tested_points_etykiety)

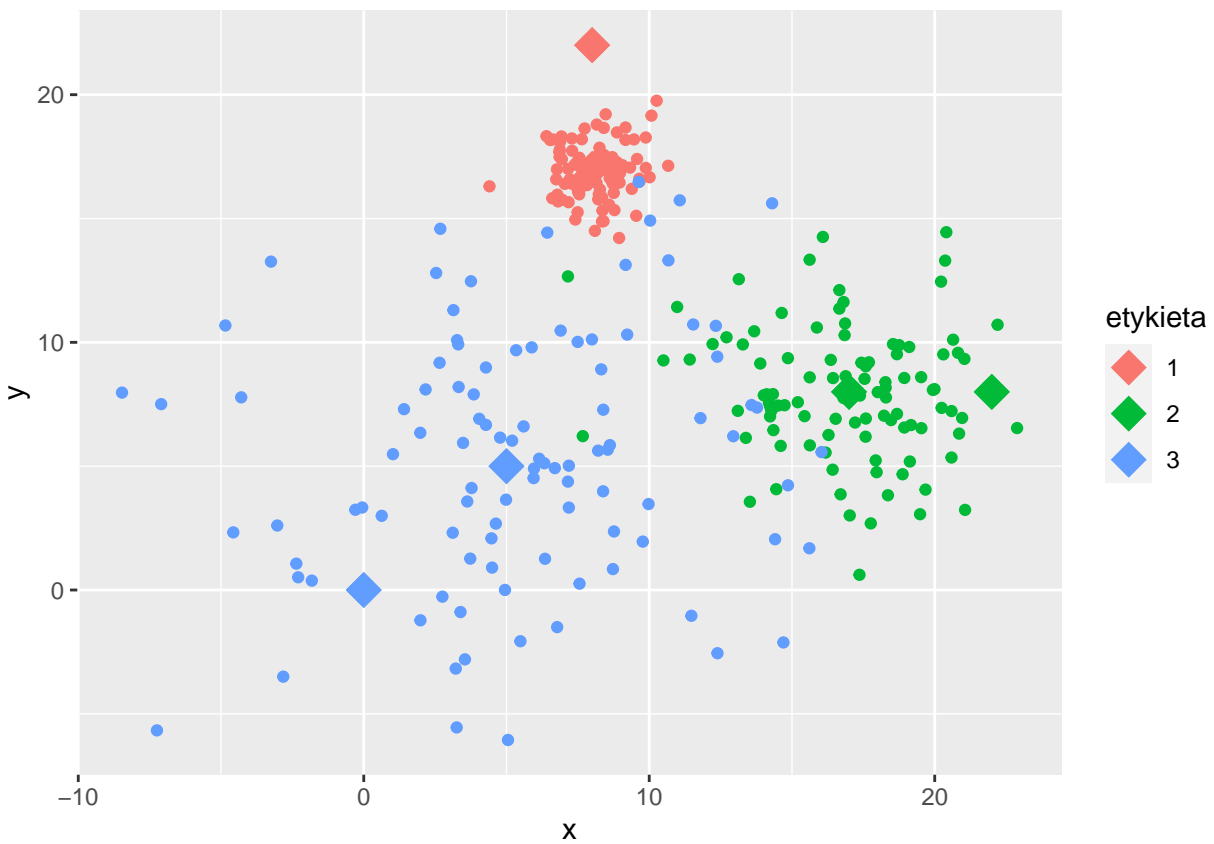
## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```

### Trzeci zbiór testowy

#### Funkcja agregująca moda

```
test_fun(third_set,third_set_etykiety,tested_points,19,2,moda,tested_points_etykiety,TRUE)
```

```
## [1] "MAD:"  
## [1] 0  
## [1] "MSE:"  
## [1] 0  
## [1] "ERR:"  
## [1] 0
```



```
test_fun(third_set,third_set_etykiety,tested_points,19,1,moda,tested_points_etykiety)
```

```
## [1] "MAD:"  
## [1] 0  
## [1] "MSE:"  
## [1] 0  
## [1] "ERR:"  
## [1] 0
```

```
test_fun(third_set,third_set_etykiety,tested_points,19,Inf,moda,tested_points_etykiety)
```

```
## [1] "MAD:"  
## [1] 0  
## [1] "MSE:"  
## [1] 0
```

```
## [1] "ERR:"  
## [1] 0
```

#### Funkcja agregująca srednia\_a

```
test_fun(third_set,third_set_etykiety,tested_points,19,2,srednia_a,tested_points_etykiety)  
  
## [1] "MAD:"  
## [1] 0  
## [1] "MSE:"  
## [1] 0  
## [1] "ERR:"  
## [1] 0  
  
test_fun(third_set,third_set_etykiety,tested_points,19,1,srednia_a,tested_points_etykiety)  
  
## [1] "MAD:"  
## [1] 0  
## [1] "MSE:"  
## [1] 0  
## [1] "ERR:"  
## [1] 0  
  
test_fun(third_set,third_set_etykiety,tested_points,19,Inf,srednia_a,tested_points_etykiety)  
  
## [1] "MAD:"  
## [1] 0  
## [1] "MSE:"  
## [1] 0  
## [1] "ERR:"  
## [1] 0
```

#### Funkcja agregująca minkara1.5

```
test_fun(third_set,third_set_etykiety,tested_points,19,2,minkara1.5,tested_points_etykiety)  
  
## [1] "MAD:"  
## [1] 0  
## [1] "MSE:"  
## [1] 0  
## [1] "ERR:"  
## [1] 0  
  
test_fun(third_set,third_set_etykiety,tested_points,19,1,minkara1.5,tested_points_etykiety)  
  
## [1] "MAD:"  
## [1] 0  
## [1] "MSE:"  
## [1] 0  
## [1] "ERR:"  
## [1] 0  
  
test_fun(third_set,third_set_etykiety,tested_points,19,Inf,minkara1.5,tested_points_etykiety)  
  
## [1] "MAD:"  
## [1] 0  
## [1] "MSE:"
```

```
## [1] 0
## [1] "ERR:"
## [1] 0
```

### Funkcja agregująca minkara3.0

```
test_fun(third_set,third_set_etykiety,tested_points,19,2,minkara3.0,tested_points_etykiety)
```

```
## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```

```
test_fun(third_set,third_set_etykiety,tested_points,19,1,minkara3.0,tested_points_etykiety)
```

```
## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```

```
test_fun(third_set,third_set_etykiety,tested_points,19,Inf,minkara3.0,tested_points_etykiety)
```

```
## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```

### Funkcja agregująca srednia\_wazona

```
test_fun(third_set,third_set_etykiety,tested_points,19,2,srednia_wazona,tested_points_etykiety)
```

```
## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```

```
test_fun(third_set,third_set_etykiety,tested_points,19,1,srednia_wazona,tested_points_etykiety)
```

```
## [1] "MAD:"
## [1] 0
## [1] "MSE:"
## [1] 0
## [1] "ERR:"
## [1] 0
```

```
test_fun(third_set,third_set_etykiety,tested_points,19,Inf,srednia_wazona,tested_points_etykiety)
```

```
## [1] "MAD:"
## [1] 0
```



```
## [1] "MSE:"  
## [1] 0  
## [1] "ERR:"  
## [1] 0
```