



ИТОГОВЫЙ ПРОЕКТ

Подготовила: Козлова Алина



ФАЙЛ С
ДАНЫМИ О
ПЛОТНОСТИ
НАСЕЛЕНИЯ И
ПРЕСТУПНОСТИ
В
МЕГАПОЛИСАХ
США

```
14  
15 data = read.table(file = "crime.txt", header =TRUE)  
16 data  
17
```

25:1 (Top Level) ↕

R Scrip

Console

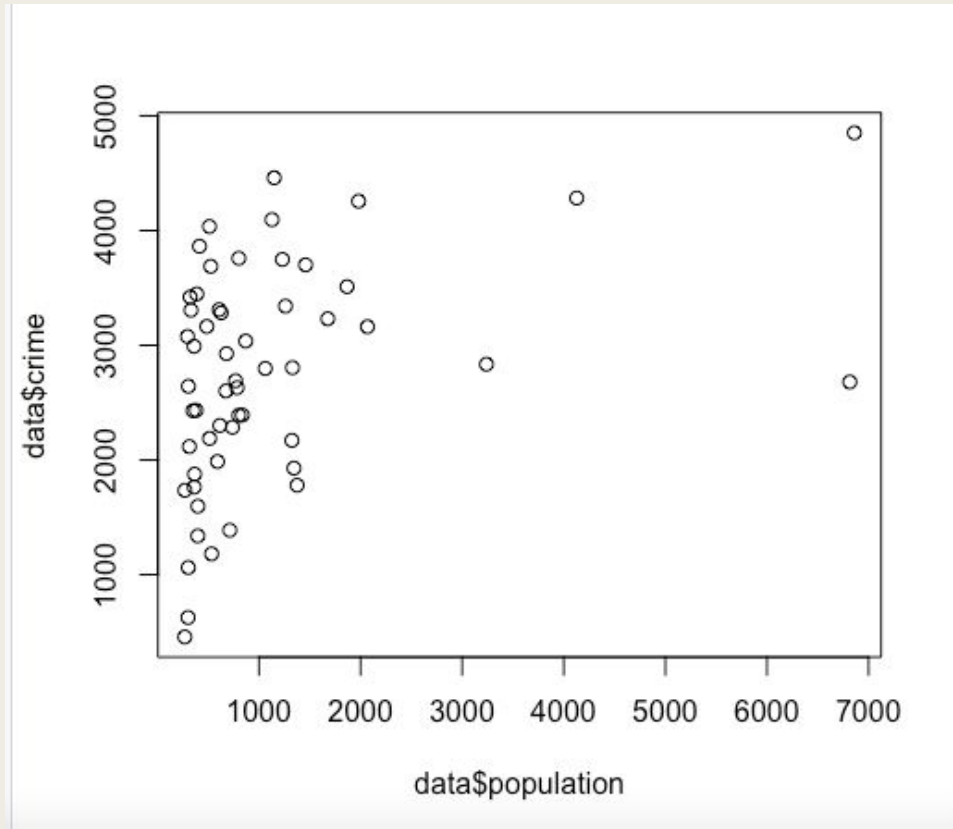
Terminal ×

Jobs ×

~/ ↻

	city	population	nonwhite	density	crime
1	Akron	675	7.3	746	2602
2	Albany	713	2.6	322	1388
3	Allentown	534	0.8	491	1182
4	Anaheim	1261	1.4	1612	3341
5	Atlanta	1330	22.8	770	2805
6	Bakersfield	331	7.0	41	3306
7	Baltimore	1981	21.6	877	4256
8	Beaumont	315	20.7	240	2117
9	Binghamton	305	0.6	147	1063
10	Birmingham	739	32.1	272	2285
11	Boston	3239	3.4	1831	2835
12	Bridgeport	785	5.3	1252	2630
13	Buffalo	1324	6.8	832	2171
14	Canton	363	5.4	630	1764
15	Charlotte	388	24.2	328	3447
16	Chattanooga	306	17.6	308	2643
17	Chicago	6815	14.8	1832	2680
18	Cincinnati	1376	10.4	640	1780

Взаимосвязь населения и преступлений



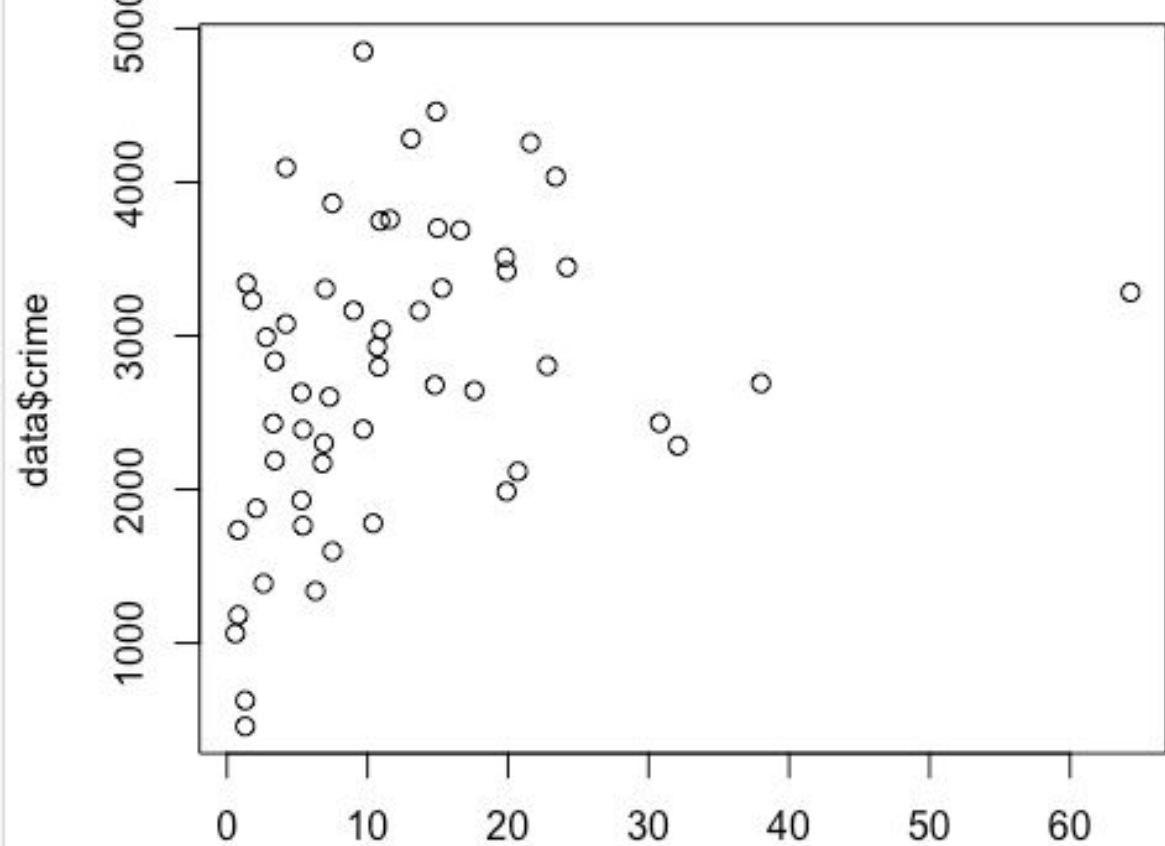
```
17  
18 cor(data$population, data$crime)  
19 plot(data$population, data$crime)|  
20
```

19:34 (Top Level) ⇅

Console Terminal × Jobs ×

~/ ↗

```
> cor(data$population, data$crime)  
[1] 0.3816201
```



```
cor(data$nonwhite, data$crime)  
plot(data$nonwhite, data$crime)
```

(Top Level) ▾

Terminal ×

Jobs ×

```
> cor(data$nonwhite, data$crime)
```

```
[1] 0.3086026
```

```
data$nonwhite, data$crime)
```

ВЗАИМОСВЯЗЬ ПРОЦЕНТА
ТЕМНОКОЖИХ ЛЮДЕЙ И
ПРЕСТУПЛЕНИЙ

```
24 summary(data)|
25
24:14 (Top Level) ↕ R Script ↕
Console Terminal × Jobs ×
~/
[1] 0.3086026
> plot(data$nonwhite, data$crime)
> summary(data)
  city          population      nonwhite
Length:55      Min.   : 270      Min.   : 0.60
Class :character 1st Qu.: 385      1st Qu.: 4.20
Mode  :character Median : 675      Median : 9.70
              Mean  :1093      Mean  :11.95
              3rd Qu.:1246      3rd Qu.:15.95
              Max.   :6860      Max.   :64.30

 density          crime
Min.   : 37.0      Min.   : 458
1st Qu.: 276.5      1st Qu.:2144
Median : 431.0      Median :2798
Mean   : 830.2      Mean   :2747
3rd Qu.: 814.0      3rd Qu.:3380
Max.   :13087.0     Max.   :4852
```

ОБЩАЯ
ИНФОРМАЦИ
Я ПО
ПАРАМЕТРАМ

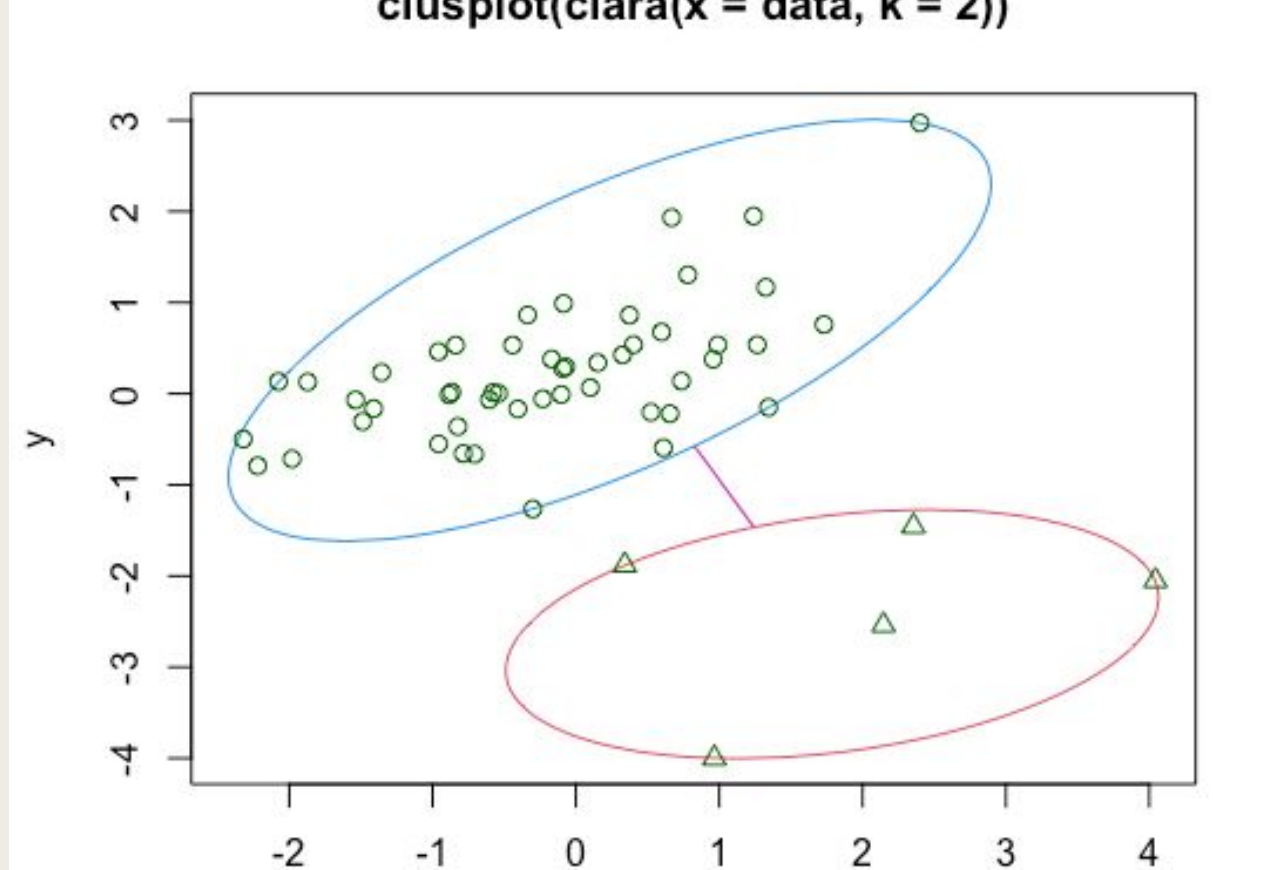
```
28 cl2 = clara(data,2)
29 cl2
30 summary(cl2)
31 #> summary
30:1 (Top Level) R Sc

Console Terminal x Jobs x

~/
Call: clara(x = data, k = 2)
Medoids:
  city population nonwhite density crime
[1,] 51          770      38.0    565 2691
[2,] 26          4127     13.1    2114 4283
Objective function: 1256.115
Clustering vector: int [1:55] 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 2 1 ...
Cluster sizes:      50 5
Best sample:
 [1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 17 18 19 21 22 23
[21] 24 25 26 27 28 29 30 31 32 33 34 35 39 40 42 43 44 45 46 49
[41] 51 52 53 54

Available components:
 [1] "sample"      "medoids"      "i.med"        "clustering"
 [5] "objective"    "clusinfo"     "diss"         "call"
 [9] "silinfo"     "data"

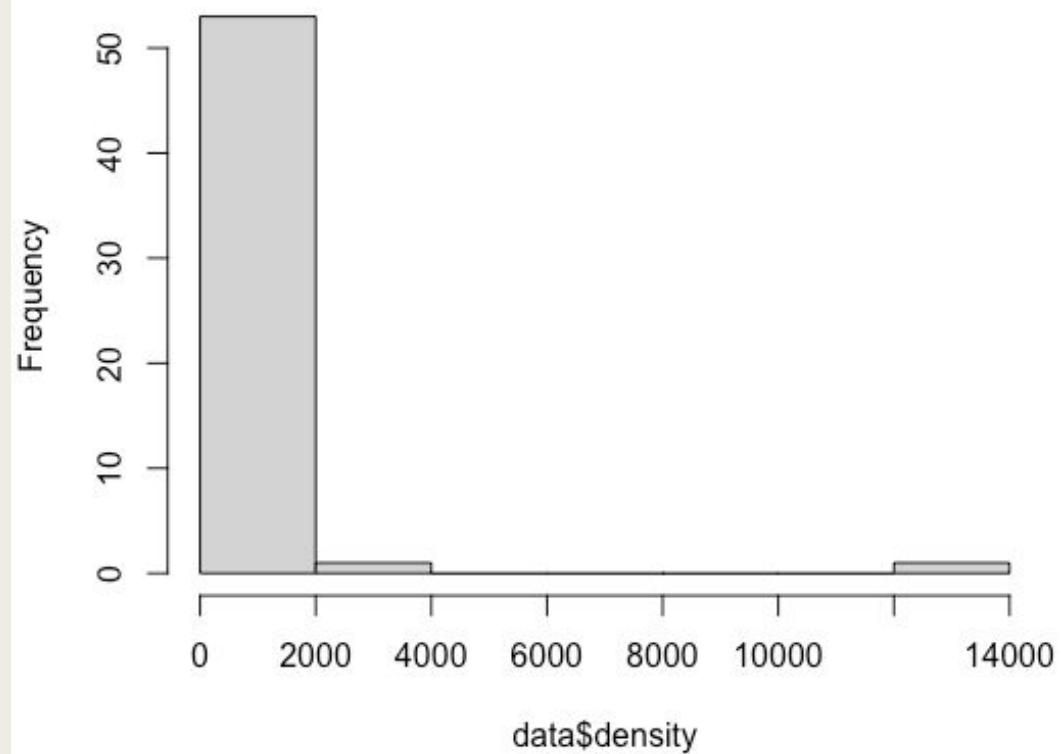
> |
```



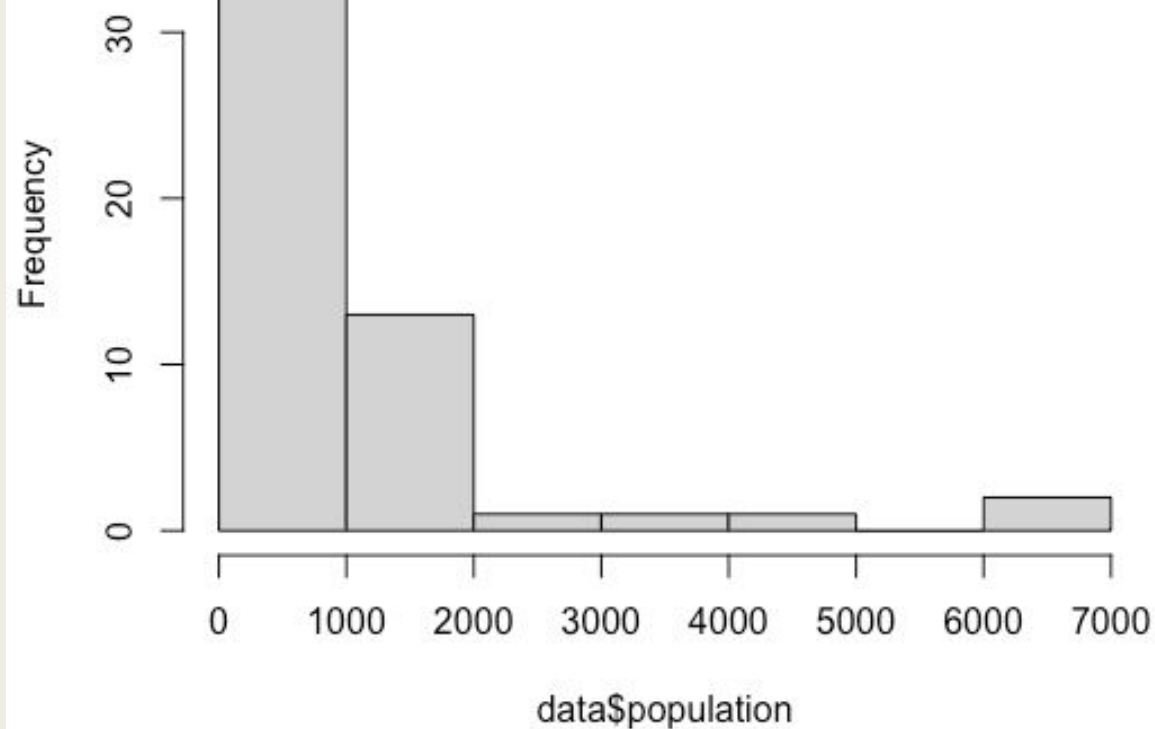
РАЗБИЕНИЕ НА КЛАСТЕРЫ

Вариационные ряды

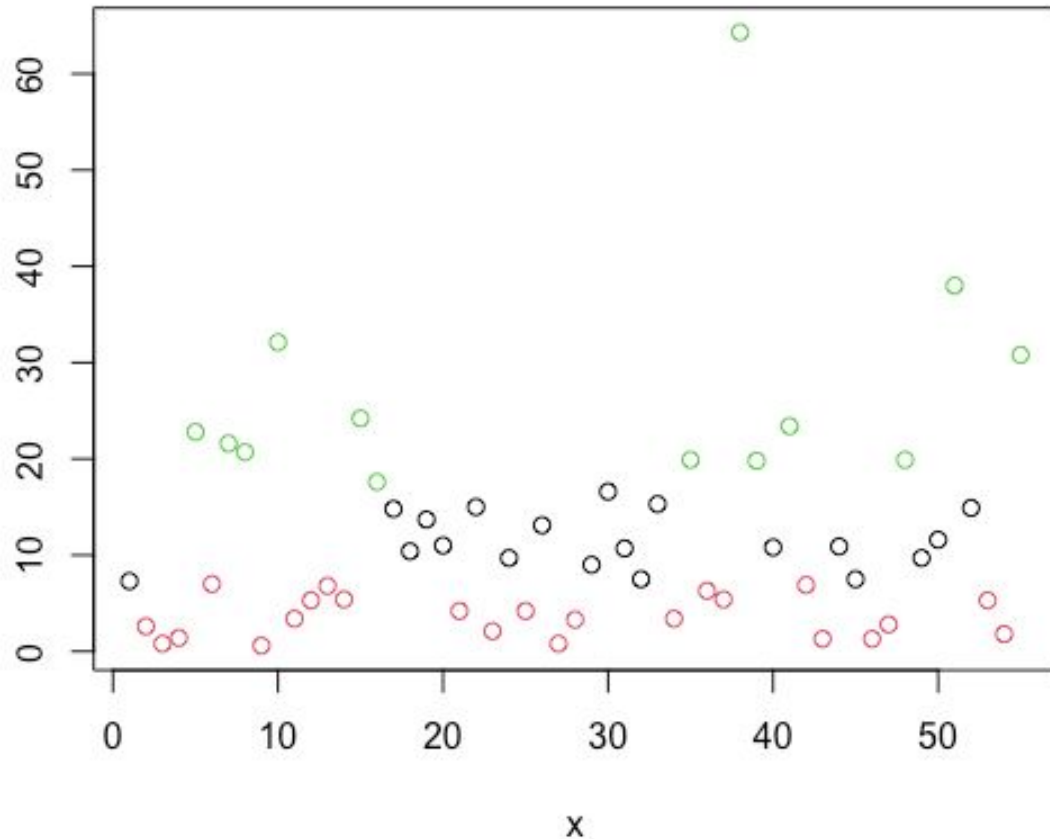
Histogram of data\$density



Histogram of data\$population



Разбиение на кластеры по параметру nonwhite



Разбиение на кластеры по преступлениям

```
dacla = data$crime  
cl7 = clara(dacla,3)  
plot(dacla, col=cl7$clustering, xlab = "x", ylab="y")
```

