

# ТЕОРИЈА УЗОРАКА час 3

19. март '14.

- Визуализација података и графици (*data visualization and graphics*)

Графичке могућности су важна и изузетно разноврсна компонента R-овог окружења. Обезбеђене су функције за генерисање различитих статистичких графика од података, као и флексибилно окружење за модификовање детаља на графицима. Могу се “градити” и потпуно нове врсте графика, прилагођене конкретном проблему.

Брз увид у обиље R-ових графика стиче се покретањем програма за демонстрацију:

```
> demo(graphics)
> demo(image)
> demo(lattice)
```

Наредбе за креирање графова деле се у три основне групе:

- функције за графике **високог нивоа** креирају нови графикон, могуће са осама, ознакама, насловима итд.
- функције за графике **ниског нивоа** додају више информација на постојећи графикон, нпр. додатне тачке, линије и ознаке
- функције за **интерактивне** графике омогућавају да се интерактивно дода информација на постојећи графикон или из њега “извуче” информација, нпр. помоћу компјутерског миша

- `plot()` - једна је од R-ових најчешће коришћених функција за графике; спада међу генеричке функције, тј. врста графика који креира зависи од врсте/класе првог аргумента.

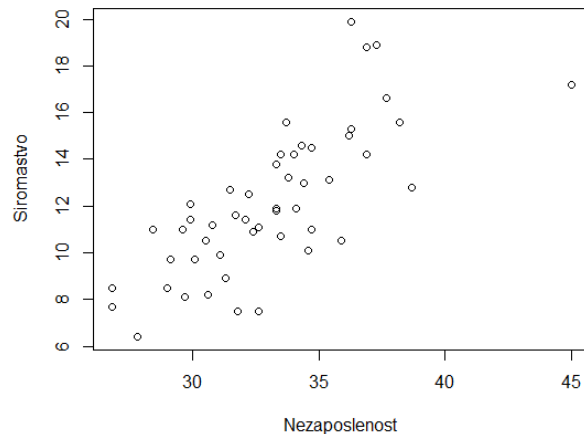
- `plot(x, y); plot(xy)`

Ако су `x` и `y` вектори, `plot(x, y)` за резултат има тачкасти график (*scatterplot*) `y`-а у зависности од `x`-а. Исти учинак може се произвести прослеђивањем само једног аргумента `xy` функцији `plot()`, који је, у том случају, листа са два вектора или матрица са две колоне.

```
> usa.data <- read.table("C:/Users/Lenchy/Desktop/posao_fax/TU/States03.csv", sep=";", header=T)
> usa.data[1:3,]
```

	State	Region	Pop	Births	Deaths	Pop18	Pop65	HSGrad	ColGrad	TeacherPay	InfMortality
1	Alabama	South	4500.752	59797	47209	25.3	13.0	79.9	22.7	38246	9.4
2	Alaska	West	648.818	9755	3133	30.4	5.7	90.6	24.0	49685	8.1
3	Arizona	West	5580.811	88154	44124	26.6	13.0	83.8	26.0	40894	6.9
	TeenBirths	ViolentCrime	PropertyCrime	DeathPenalty	Executions	Poverty	Unemp	Uninsured			
1	54.5	444.2	4020.9	Yes	28	16.6	37.7	13.3			
2	39.5	563.4	3746.3	No	0	7.7	26.8	17.8			
3	61.2	552.9	5833.4	Yes	22	14.2	36.9	17.3			
	Income	Earnings	Heart	Vehicles	Homeowners						
1	35412	13.56	289.3	23.1	76.2						
2	56536	12.16	187.7	17.4	70.0						
3	41172	14.38	204.5	19.4	67.0						

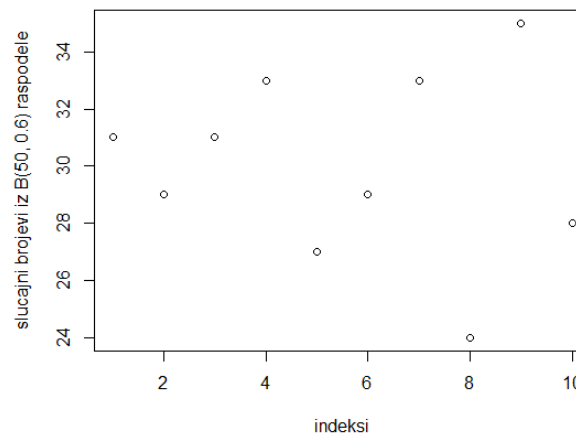
```
> plot(usa.data$Unemp, usa.data$Poverty, xlab="Nezaposlenost", ylab="Siromastvo")
> plot(cbind(usa.data$Unemp, usa.data$Poverty), xlab="Nezaposlenost", ylab="Siromastvo")
```



- `plot(x)`

Ako je  $x$  vremenska serija (*time series*), rezultat je graф временске серије; ако је  $x$  нумерички вектор на излазу се добија график вредности тог вектора у зависности од индекса у вектору.

- > `plot(airmiles, main="Podaci o vazdusnim miljama", xlab="vreme", ylab="Prihodi putnicke milje na komercijalnim letovima u SAD")`
- > `x <- rbinom(10, 50, .6)`
- > `plot(x, xlab="indeksi", ylab="slucajni brojevi iz B(50, 0.6) raspodele")`



- `plot(f)`; `plot(f, y)`

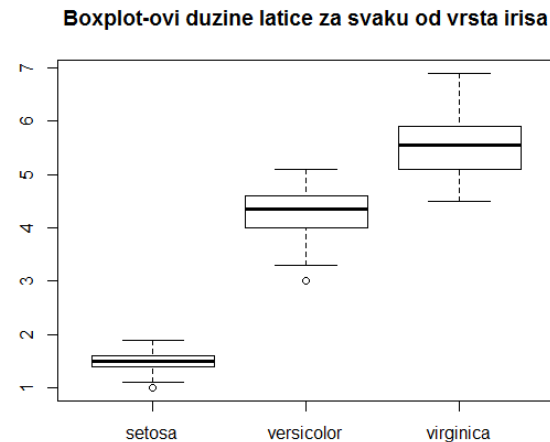
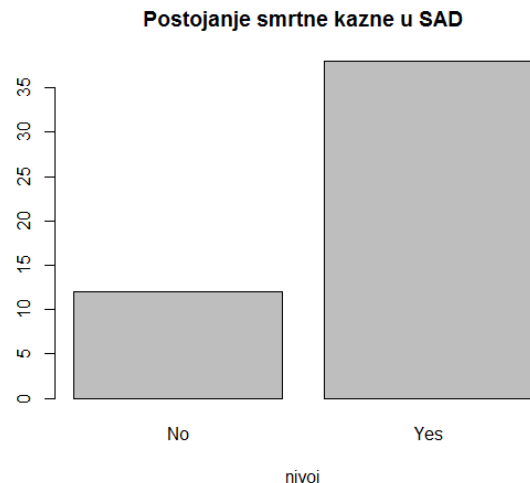
`f` је фактор, `y` је нумерички вектор; првом наредбом генерише се график `f`-а у облику *bar*-ова, док се као резултат друге наредбе добијају *boxplot*-ови `y`-а за сваки ниво `f`-а.

```
> class(usa.data$DeathPenalty)
```

```
[1] "factor"
```

```
> plot(usa.data$DeathPenalty, xlab="nivoi", main="Postojanje smrtne kazne u SAD")
```

```
> plot(iris$Species, iris$Petal.Length,
      main="Boxplot-ovi duzine latice za svaku od vrsta irisa")
```



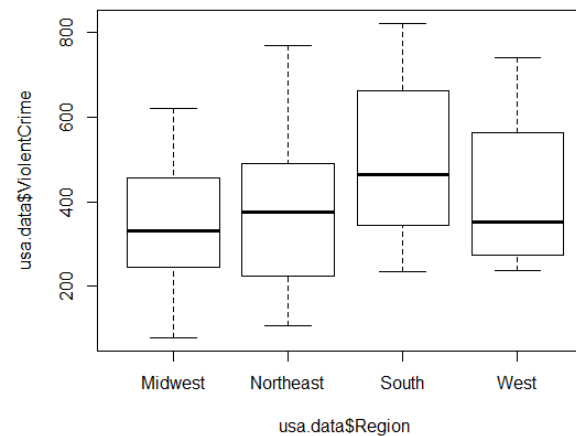
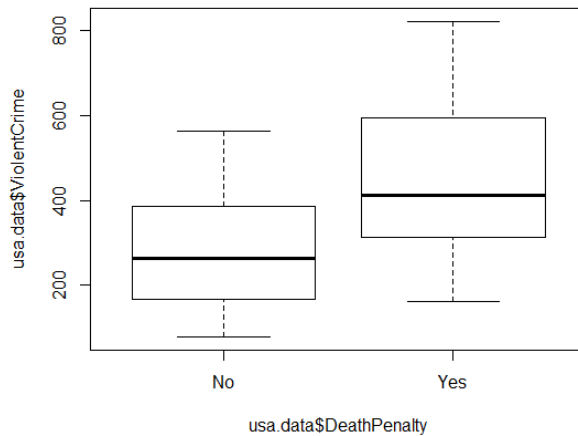
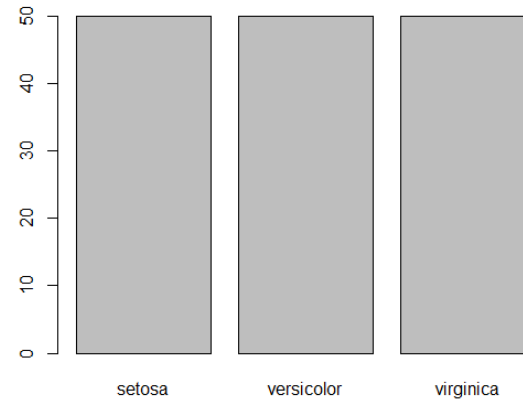
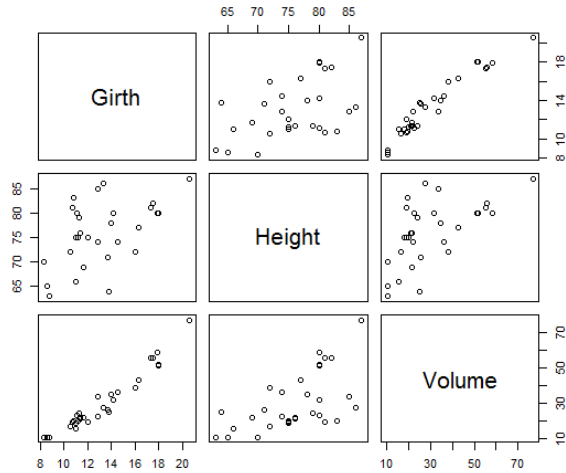
- `plot(df)`; `plot(~ expr)`; `plot(y ~ expr)`

`df` је база података, `y` је било који објекат, `expr` је списак назива објеката одвојених помоћу '+'; првим двама наредбама, редом, генеришу се *distribution plots* променљивих `y` у бази, односно *distribution plots* броја именованих објеката; излаз треће наредбе је низ графика `y`-а у односу на сваки објекат именован у `expr`.

```

> plot(trees)
> plot(~ iris$Species)
> plot(usa.data$ViolentCrime ~ usa.data$DeathPenalty + usa.data$Region)
Hit <Return> to see next plot:
Hit <Return> to see next plot:
> #isprobati i: plot(usa.data$ViolentCrime ~ usa.data$DeathPenalty + usa.data$Unemp)

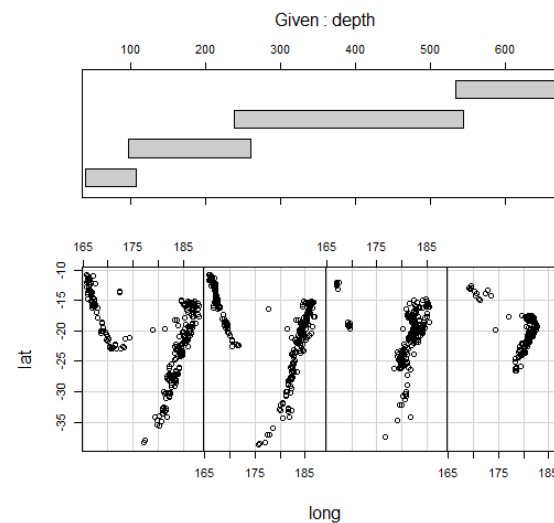
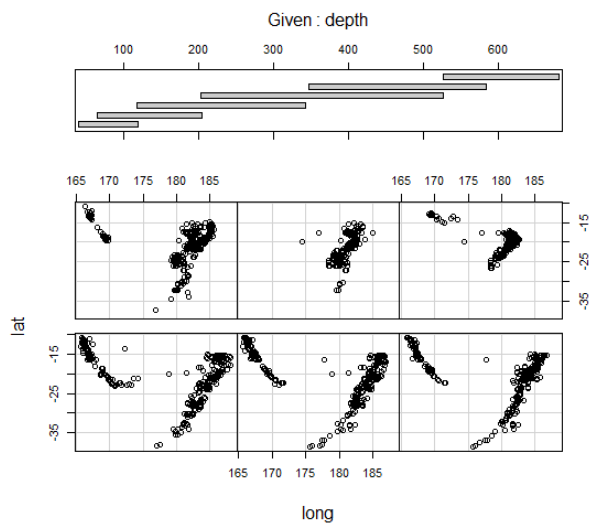
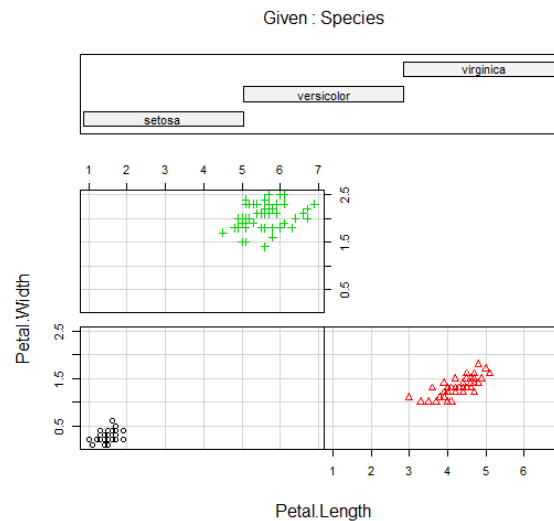
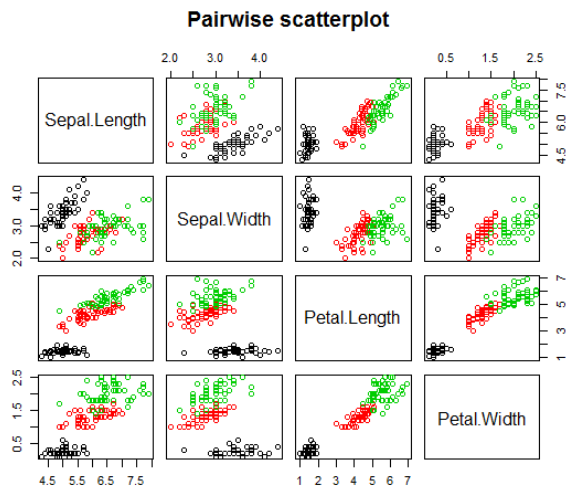
```



У R постоје и две веома корисне функције за презентовање вишеструких података.

- `pairs(X)`, где је `X` нумеричка матрица или база података - је наредба којом се добија матрица тачкастих графика парова променљивих дефинисаних колонама `X`-а, тј. свака колона `X`-а приказана је у зависности од сваке друге колоне `X`-а.
- `coplot(a ~ b | c)`, где су `a` и `b` нумерички вектори, а `c` вектор или фактор (претпоставља се да су сви једнаке дужине) - је наредба чији је излаз низ графика, на којима је приказана зависност `a` од `b` за дате вредности `c`-а. Ако је `c` фактор, то значи да је `a` приказан у односу на `b` за сваки ниво `c`-а. Ако је, пак, `c` нумерички вектор, он је издељен на неколико “условних” интервала и за сваки интервал `a` је приказан у односу на `b` за вредности `c`-а унутар тог интервала.

```
> attach(iris)
> pairs(iris[,1:4], col=as.numeric(Species), main="Pairwise scatterplot")
> coplot(Petal.width ~ Petal.Length | Species, col=as.numeric(Species), pch=as.numeric(Species))
> coplot(lat ~ long | depth, data = quakes)
> given.depth <- co.intervals(quakes$depth, number = 4, overlap = .1)
> coplot(lat ~ long | depth, data = quakes, given.v = given.depth, rows = 1)
```



Обе функције имају опционални аргумент `panel`, који је намењен прилагођавању врсте графика у сваком панелу.



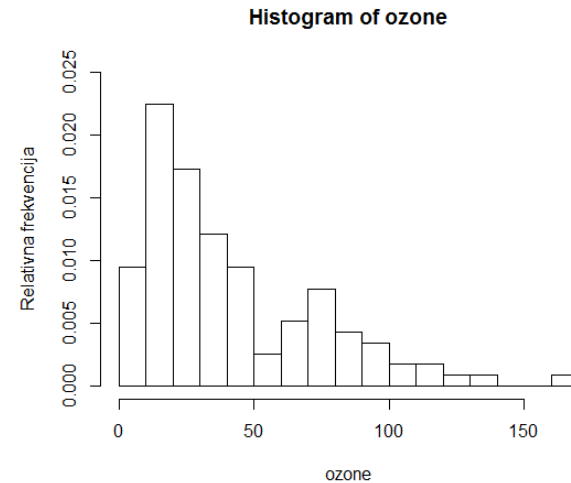
- `hist(x)`, где је `x` нумерички вектор - је наредба за исцртавање хистограма (то је посебна врста графика са *bar*-овима (*bin*-овима), која се користи за приказ расподеле фреквенција низа (непрекидних) нумеричких вредности). Од опционалних аргумената важно је издвојити: `nclass` - постављањем на одређену вредност корисник може да препоручи жељени (разуман) број *bar*-ова; алтернативно, број *bar*-ова, односно граничне тачке *bar*-ова могу се тачно задати аргументом `breaks` (`breaks="sturges"` / `breaks="scott"` / `breaks="Freedman-Diaconis"`); ако је наведено `prob=T`, вредности на у-оси представљају релативне, а не апсолутне фреквенције.

```
> ozone <- airquality$ozone
```

```
> hist(ozone)
```

```
> hist(ozone, breaks=15, prob=T, ylim=c(0, 0.025), ylab="Relativna frekvencija")
```

Хистограм је најједноставнија непараметарска “оцена” густине расподеле. Преферирана алтернатива хистограму је исцртавање “оцене” густине расподеле базиране на језгру (*kernel density plot*).

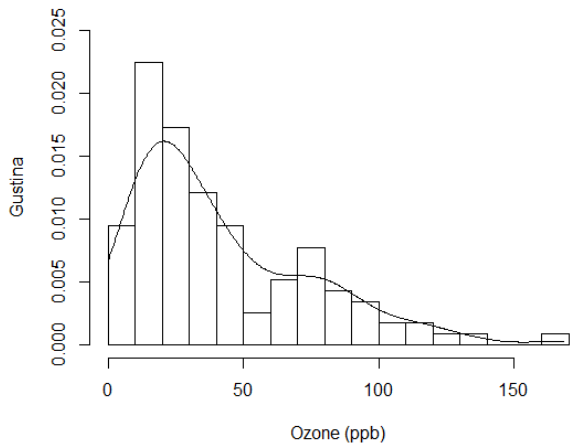


```

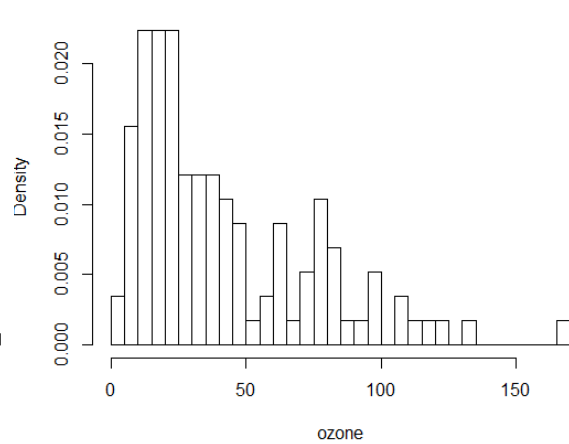
> max.ozone <- max(ozone, na.rm=T)
> hist(ozone, breaks=15, freq=F, xlab='ozone (ppb)', ylim=c(0, 0.025),
      ylab='Gustina', main='Histogram sa grafikom gustine jezgra')
> lines(density(ozone, na.rm=T, from=0, to=max.ozone))
> mean.ozone <- mean(ozone, na.rm=T); sd.ozone <- sd(ozone, na.rm=T)
> ozone.histogram <- hist(ozone, breaks=50, freq=F)
> ozone.ylim.normal <- range(0, ozone.histogram$density, dnorm(ozone,
      mean=mean.ozone, sd=sd.ozone), na.rm=T)
> hist(ozone, breaks=15, freq=F, ylim=c(0, 0.025), xlab='Ozone (ppb)',
      ylab='Gustina', main='Histogram sa grafikom gustine normalne raspodele')
> curve(dnorm(x, mean=mean.ozone, sd=sd.ozone), add = T)

```

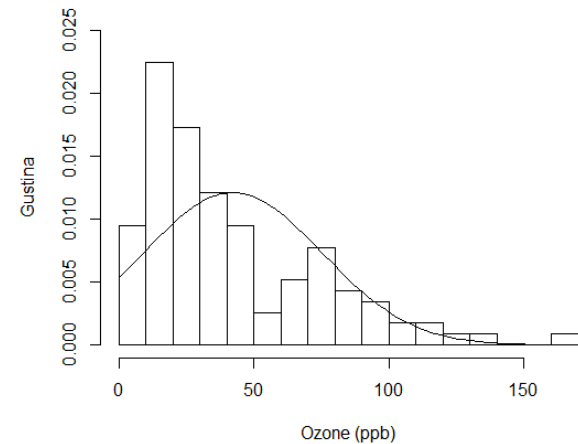
Histogram sa grafikom gustine jezgra



Histogram of ozone



Histogram sa grafikom gustine normalne raspodele



```

> var.ozone <- var(ozone, na.rm=T)
> #isprobati i:hist(ozone, breaks=15, freq=F, xlim=c(0, 170), ylim=c(0, 0.025), xlab='Ozone
(ppb)', ylab = 'Relativna frekvencija', main = 'Histogram sa krivom gustine gama raspodele')
> #curve(dgamma(x, shape = mean.ozone^2/var.ozone, scale = var.ozone/mean.ozone), add = T)

```

Иначе, многи методи за генерисање графика високог нивоа враћају резултат, који се може доделити објекту или се користити директно у изразима. Нпр. функција/метод `hist(x)` враћа граничне и средишње тачке *bar*-ова, апсолутне и релативне фреквенције:

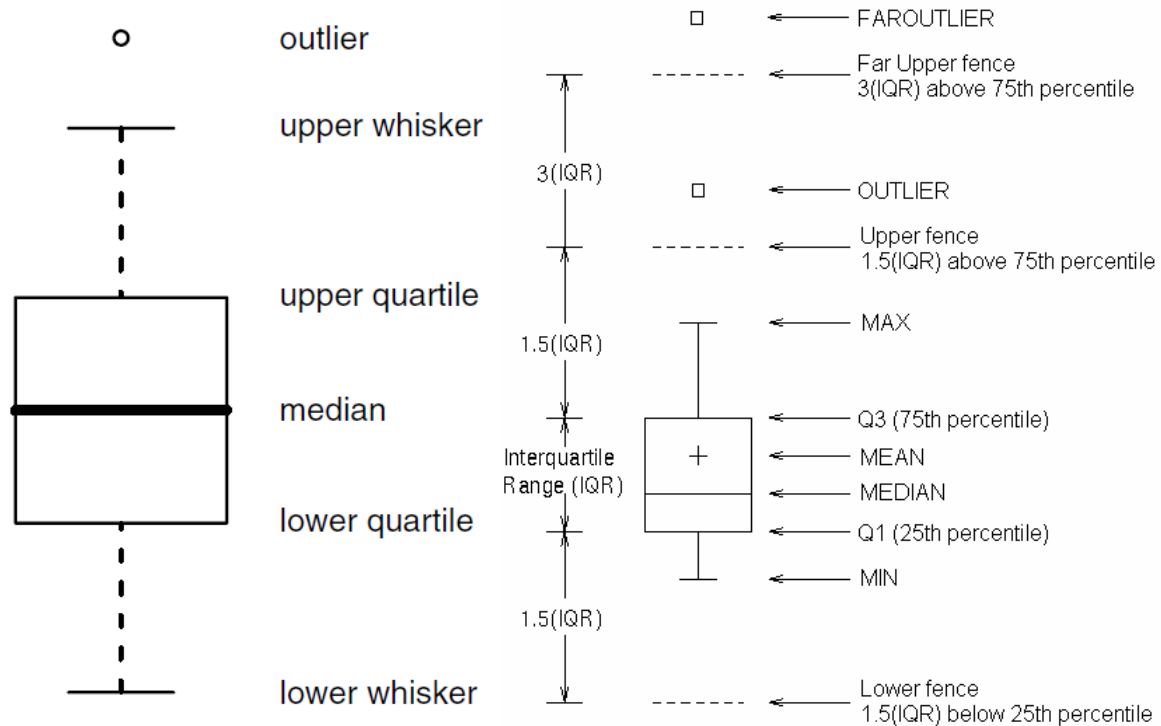
```

> str(ozone.histogram)
List of 6
 $ breaks   : num [1:18] 0 10 20 30 40 50 60 70 80 90 ...
 $ counts   : int [1:17] 11 26 20 14 11 3 6 9 5 4 ...
 $ density  : num [1:17] 0.00948 0.02241 0.01724 0.01207 0.00948 ...
 $ mids     : num [1:17] 5 15 25 35 45 55 65 75 85 95 ...
 $ xname    : chr "ozone"
 $ equidist: logi TRUE
 - attr(*, "class")= chr "histogram"

```

- `boxplot(x)`, где је `x` нумерички вектор - је наредба за исцртавање правоугаоног дијаграма. Приликом статистичке анализе требало би га обавезно користити за приказивање (непрекидне) нумеричке променљиве, чије су вредности измерене на подгрупама субјеката, одређеним категоријском променљивом.

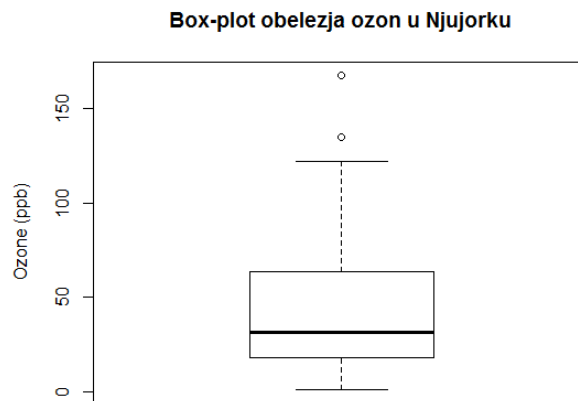
Висина правоугаоника (*box-a*) дефинисана је распоном између квантила (IQR). *The lower whisker* (доња “ограда”) повлачи се са доње стране *box-a* и налази се на нивоу који одговара већој од две вредности: узорачки минимум и први квантил умањен за 1.5IQR. Слично, *the upper whisker* (горња “ограда”) повлачи се са горње стране *box-a* и налази се на нивоу који одговара мањој од две вредности: узорачки максимум и трећи квантил увећан за 1.5IQR.



> summary(ozone)

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
1.00	18.00	31.50	42.13	63.25	168.00	37

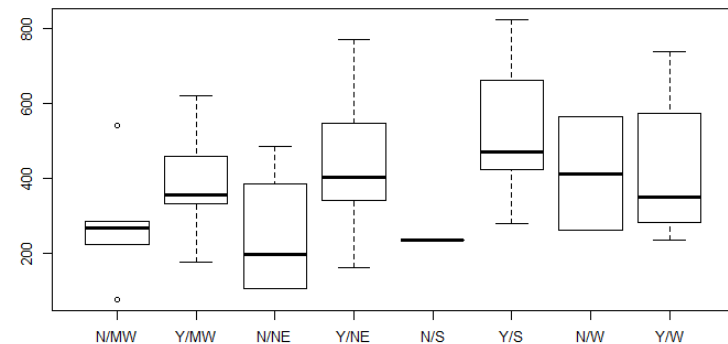
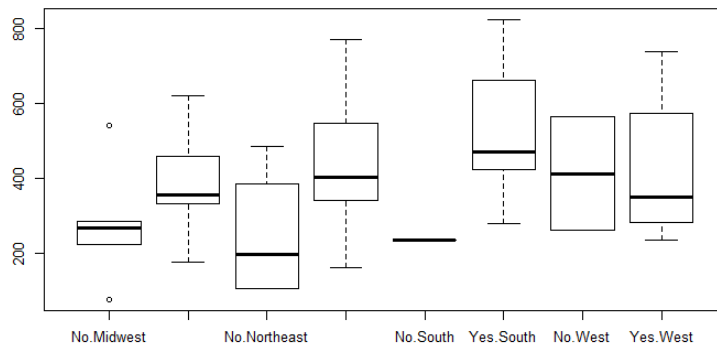
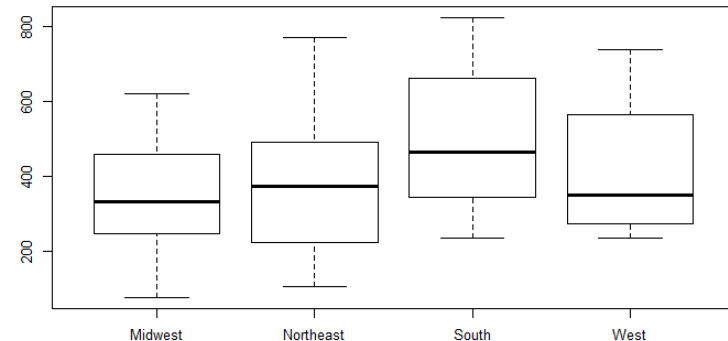
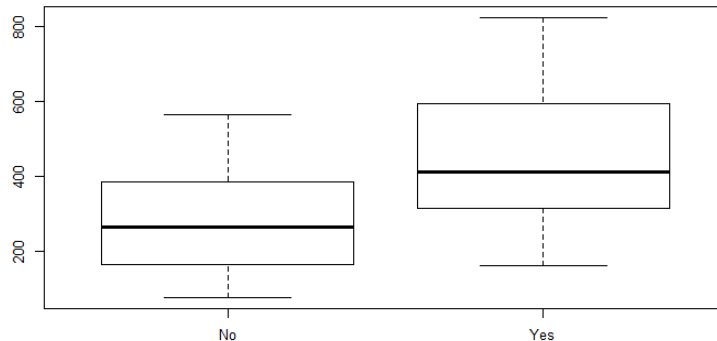
> boxplot(ozone, ylab = 'Ozone (ppb)', main = 'Box-plot obelezja ozon u Njujorku')



```

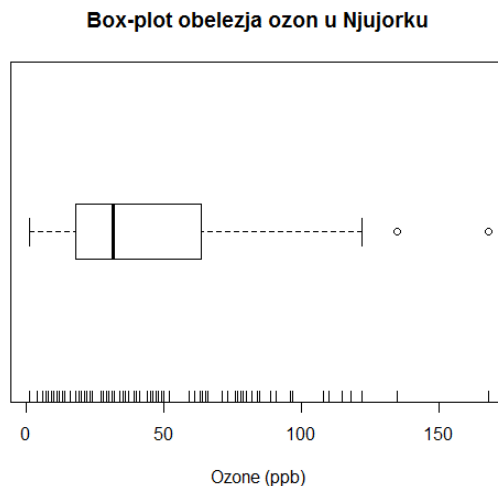
> par(mfrow=c(2, 2), mar=c(3, 3, 2, 1))
> boxplot(ViolentCrime ~ DeathPenalty, data=usa.data)
> boxplot(ViolentCrime ~ Region, data=usa.data)
> boxplot(ViolentCrime ~ DeathPenalty * Region, data=usa.data)
> boxplot(ViolentCrime ~ DeathPenalty * Region, data=usa.data,
          names=c("N/MW", "Y/MW", "N/NE", "Y/NE", "N/S", "Y/S", "N/W", "Y/W"))

```



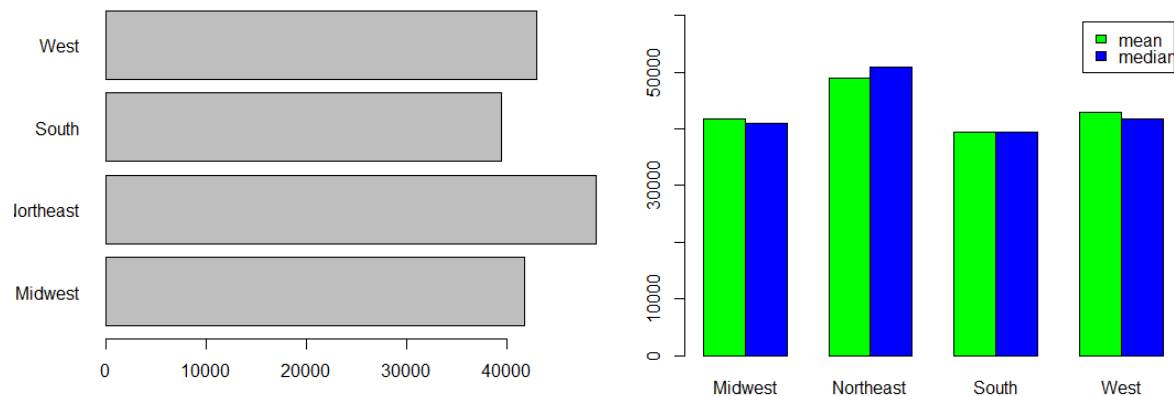
Уместо вертикалне оријентације *boxplot*-ова, навођењем опционалног аргумента `horizontal=T`, добијају се хоризонтално постављени *boxplot*-ови.

```
> boxplot(ozone, xlab = 'Ozone (ppb)', boxwex=.35, horizontal=T,  
          main = 'Box-plot obelezja ozon u Njujorku')  
> rug(ozone, side=1)
```



- `barplot()` - je funkcija za generisanje *bar*-diјаграма. Као аргумент прослеђује јој се објекат који чува податке, а обично је то вектор или матрица (табела).

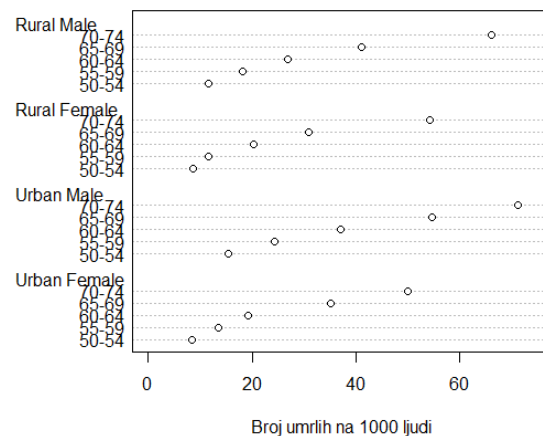
```
> m1 <- tapply(usa.data$TeacherPay, usa.data$Region, mean)  
> barplot(m1, horiz=T, las=1)  
> m2 <- tapply(usa.data$TeacherPay, usa.data$Region, median)  
> r <- rbind(m1, m2)  
> b <- barplot(r, col=c("green", "blue"), ylim=c(0, 60000), beside=T,  
              legend=c("mean", "median"), pch=15)
```



▫ `dotchart()` - je funkcija za generisanje tачкаstog grafika. Posebno je pogodan za vizualni prikaz svih unesenih podataka s vrednostima koje leže u određenim rasponima.

> `dotchart(VADeaths, xlim=c(0,75), xlab="Broj umrlih na 1000 ljudi", main="Stopa smrtnosti u Virdziniji")`

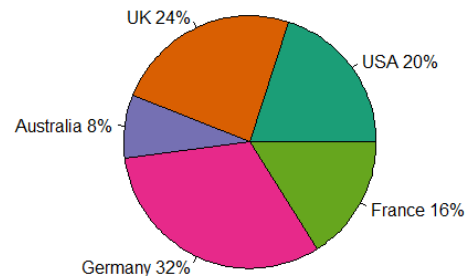
Stopa smrtnosti u Virdziniji



- `pie()` - је функција за генерисање кружног дијаграма. Као аргумент прослеђује јој се објекат који чува податке, а обично је то вектор или матрица.

```
> library("RColorBrewer", lib.loc="C:/Users/Lenchy/Documents/R/win-library/3.0")
> slices <- c(10, 12, 4, 16, 8)
> lbls <- c("USA", "UK", "Australia", "Germany", "France")
> pct <- round(slices/sum(slices)*100)
> lbls <- paste(lbls, pct); lbls <- paste(lbls, "%", sep="")
> jBrewColors <- brewer.pal(n = 5, name = "Dark2")
> pie(slices, labels=lbls, col=jBrewColors, main="Pie Chart of Countries")
```

Pie Chart of Countries



- `qqnorm(x)`, `qqline(x)`, `qqplot(x, y)` - qq-графици су врста тачкастих графика, који се користе да би се упоредиле расподеле две групе података или узорачка са (неком) референтном расподелом.

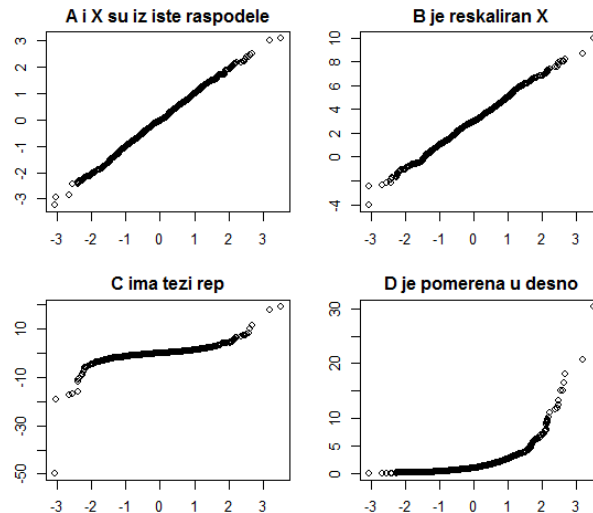
```
> #qqnorm(x) iscrtava numericki vektor x u odnosu na ocekivanu normalnu raspodelu
> #qqline(x) dodaje pravu na prethodni grafik koja prolazi kroz kvartile raspodele i podataka
> #qqplot(x, y) iscrtava kvantile x-a u odnosu na kvantile y-a radi poredjenja njihovih pripadnih raspodela
```



```

> X <- rnorm(1000); A <- rnorm(1000); B <- rnorm(1000, mean=3, sd=2)
> C <- rt(1000, df=2); D <- exp(rnorm(1000))
> par(mfrow=c(2, 2), mar=c(3, 3, 2, 1))
> qqplot(X, A, main="A i X su iz iste raspodele")
> qqplot(X, B, main="B je reskaliran X")
> qqplot(X, C, main="C ima tezi rep")
> qqplot(X, D, main="D je pomerena u desno")

```



Који тип графика би требало одабрати зависи од више фактора и то, пре свега, од:

- типа података
- намене графичког приказа
- функционисања визуалне перцепције код човека.

▫ Аргументи за функције графика високог нивоа

У питању су опционални аргументи, који су идентични за више поменутих функција за исцртавање графика:

опција	опис
<code>add=F</code> (default)	ако је постављен на <code>T</code> делује као функција за график ниског нивоа, тј. додаје нови график на већ постојећи (само неке функције)
<code>axes=T</code> (default)	ако је постављен на <code>F</code> не исцртава координатне осе
<code>type='p'</code> (default)	контролише врсту графа који се креира; могуће вредности су <code>'p'</code> , <code>'l'</code> , <code>'b'</code> , <code>'c'</code> , <code>'o'</code> , <code>'h'</code> , <code>'s'</code> , <code>'S'</code> .
<code>pch=</code>	карактер којим су приказане тачке на графику; могуће вредности су <code>1, 2, ..., 25</code> или било који карактер под наводницима
<code>cex=</code> ; <code>cex.main=</code>	подешава величину карактера којим су приказане тачке на графику; могуће вредности су <code>1, 1.5, 2, ...</code>
<code>lty=</code>	подешава тип линије; могуће вредности су <code>1, 2, ...</code>
<code>lwd=</code>	подешава дебљину линије; могуће вредности су <code>1, 2, ...</code>
<code>col=</code>	подешава боју карактера којим су приказане тачке на графику
<code>xlim=</code> ; <code>ylim=</code>	одређује доњу и горњу граничну вредност на одговарајућој оси;
<code>xlab=</code> ; <code>ylab=</code>	одређује ознаку одговарајуће осе; могућа вредност је стринг
<code>main=</code>	одређује наслов графика, смештен при врху и исписан великим фонтом; могућа вредност је стринг
<code>sub=</code>	одређује поднаслов графика, смештен при дну, испод <code>x</code> -осе и исписан мањим фонтом; могућа вредност је стринг
<code>font.axis=</code>	подешава изглед фонта; могуће вредности су <code>1, 2, 3, 4</code>

## ▫ Додавање елемената на већ постојећи график

Понекад функције за графике високог нивоа не креирају у потпуности онакав график какав корисник жели. У том случају могу се искористити функције за графике ниског нивоа, ради додавања екстра информација (нпр. тачака, линија, текста) на тренутни график. Неке од корисних функција су и:

- `points(x, y); lines(x, y)` - додају се тачке, односно линије на тренутно активан график
- `text(x, y, labels)` - додаје се текст у тачки с координатама (x, y); често се користи у низу:

```
> plot(x, y, type="n"); text(x, y, names)
```

- `abline(a, b); abline(h=y); abline(v=x); abline(lm.obj)` - доцртава се права на тренутно активни граф; првом наредбом додаје се права  $y=a+bx$ , наредним двама наредбама додаје се права паралелна x-оси, односно y-оси
- `legend(x, y, legend, ...)` - доцртава се легенда тренутно активном графику, на одређеној позицији
- `rug(x)` - доцртавају се подаци из x-a на x-оси као вертикалне цртице

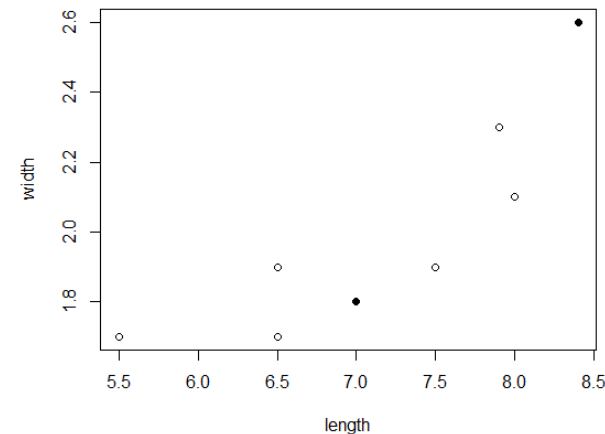
- `title(main, sub, xlab, ylab)` - додаје се главни наслов, поднаслов или ознаке оса
- `mtext(text, side, line)` - додаје се текст у маргини
- `axis(side, at, labels)` - додаје се оса тренутно активном графику; и то на страни одређеној аргументом `side`, који може бити постављен на једну од вредности 1, 2, 3, 4 (бројећи у смеру кретања казаљке на сату, и почевши од дна)
- `box()` - додаје се правоугаоник око области са графиком
- `grid()` - додаје се координатна мрежа
- `curve(...)` - додаје се крива која одговара одређеној функцији или изразу

```
> index.finger #mala baza podataka
```

	sex	length	width
1	M	7.9	2.3
2	F	6.5	1.7
3	M	8.4	2.6
4	F	5.5	1.7
5	F	6.5	1.9
6	M	8.0	2.1
7	F	7.0	1.8
8	M	7.5	1.9

```
> plot(width ~ length, data=index.finger)
```

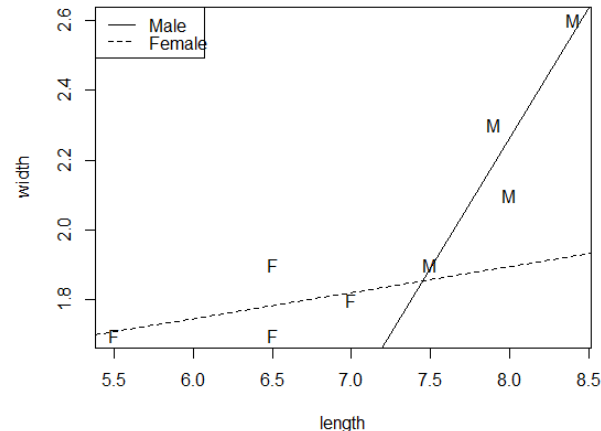
```
> with(index.finger[c(3,7),], points(length, width, pch=16))
```



```

> plot(width ~ length, pch=as.character(sex), data=index.finger)
> abline(-3.689, 0.7439, lty=1)
> abline(1.30526, .07368, lty=2)
> legend("topleft", legend=c("Male", "Female"), lty=1:2)

```



## ▫ Постављање графичких параметара

Функцијом `par()` могу се вршити подешавања читавог низа графичких параметара. За већину графика default подешавања су адекватна. Када корисник ипак жели да измени вредности одређених аргумената, требало би да пре икаквих измена (а све измене су перманентне) и исцртавања графика сачува постојећа, оригинална подешавања:

```
> oldpar <- par(no.readonly=T)
```

и да их, касније, по потреби, врати:

```
> par(oldpar)
```

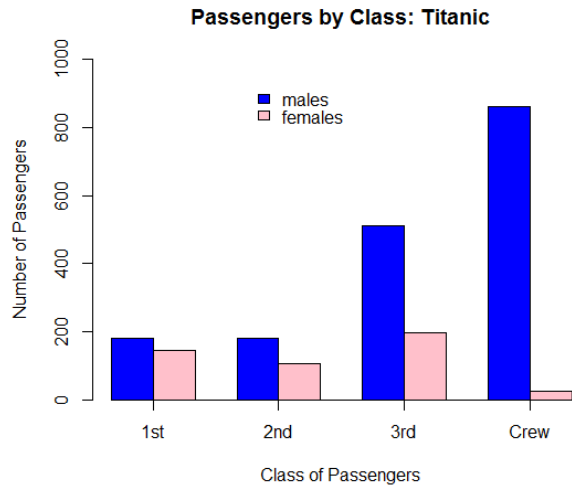
<code>adj</code>	controls text justification with respect to the left border of the text so that 0 is left-justified, 0.5 is centred, 1 is right-justified, values > 1 move the text further to the left, and negative values further to the right; if two values are given (e.g., <code>c(0, 0)</code> ) the second one controls vertical justification with respect to the text baseline
<code>bg</code>	specifies the colour of the background (e.g., <code>bg="red"</code> , <code>bg="blue"</code> ; the list of the 657 available colours is displayed with <code>colors()</code> )
<code>bty</code>	controls the type of box drawn around the plot, allowed values are: "o", "l", "7", "c", "u" ou "]" (the box looks like the corresponding character); if <code>bty="n"</code> the box is not drawn
<code>cex</code>	a value controlling the size of texts and symbols with respect to the default; the following parameters have the same control for numbers on the axes, <code>cex.axis</code> , the axis labels, <code>cex.lab</code> , the title, <code>cex.main</code> , and the sub-title, <code>cex.sub</code>
<code>col</code>	controls the colour of symbols; as for <code>cex</code> there are: <code>col.axis</code> , <code>col.lab</code> , <code>col.main</code> , <code>col.sub</code>
<code>font</code>	an integer which controls the style of text (1: normal, 2: italics, 3: bold, 4: bold italics); as for <code>cex</code> there are: <code>font.axis</code> , <code>font.lab</code> , <code>font.main</code> , <code>font.sub</code>
<code>las</code>	an integer which controls the orientation of the axis labels (0: parallel to the axes, 1: horizontal, 2: perpendicular to the axes, 3: vertical)
<code>lty</code>	controls the type of lines, can be an integer (1: solid, 2: dashed, 3: dotted, 4: dotdash, 5: longdash, 6: twodash), or a string of up to eight characters (between "0" and "9") which specifies alternatively the length, in points or pixels, of the drawn elements and the blanks, for example <code>lty="44"</code> will have the same effet than <code>lty=2</code>
<code>lwd</code>	a numeric which controls the width of lines
<code>mar</code>	a vector of 4 numeric values which control the space between the axes and the border of the graph of the form <code>c(bottom, left, top, right)</code> , the default values are <code>c(5.1, 4.1, 4.1, 2.1)</code>
<code>mfc</code>	a vector of the form <code>c(nr,nc)</code> which partitions the graphic window as a matrix of <code>nr</code> lines and <code>nc</code> columns, the plots are then drawn in columns
<code>mfrow</code>	id. but the plots are then drawn in line
<code>pch</code>	controls the type of symbol, either an integer between 1 and 25, or any single character within ""
<code>ps</code>	an integer which controls the size in points of texts and symbols

## ▫ Интерактивни рад са графицима

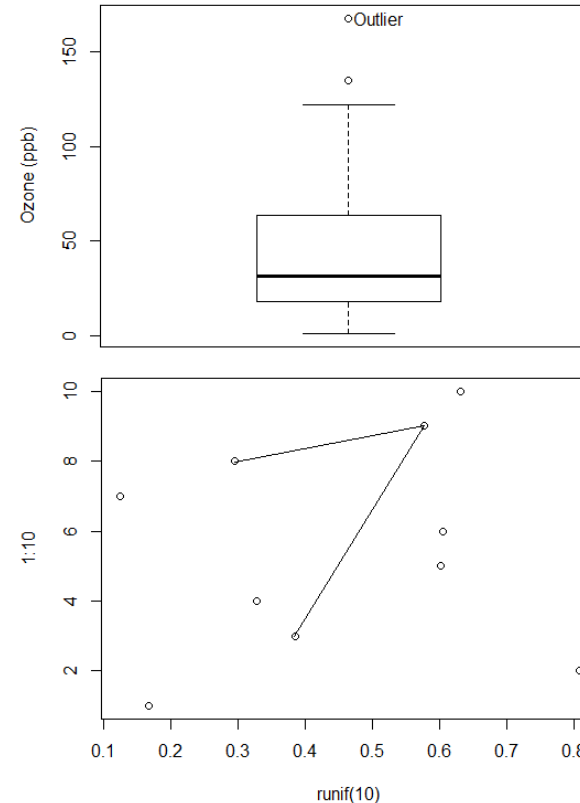
R, такође, поседује и функције које кориснику омогућавају “извлачење” или додавање информација на графику, коришћењем компјутерског миша.

- `Locator(n, type)` - чека се док корисник не одабере локације на тренутно активном графику коришћењем леве типке миша. Опционални аргумент `type` омогућава исцртавање графика од одабраних тачака и има исти учинак као и функције за графике високог нивоа. `Locator()` враћа локације одабраних тачака као листу са две компоненте `x` и `y`. Ова функција обично се позива без аргумената. Посебно је корисна за интерактивно одабирање позиција графичких елемената као што су легенде или ознаке. када је тешко унапред израчунати на које би их координате требало позиционирати.

```
> data.table <- margin.table(Titanic, c(2,1))
> barplot(data.table, beside=T, axis.lty=1, ylim=c(0,1000), col=c("blue","pink"))
> title(xlab="Class of Passengers", ylab="Number of Passengers")
> title(main="Passengers by Class: Titanic")
> legend(locator(1), legend=c("males","females"), fill=c("blue","pink"), bty="n")
> boxplot(ozone, ylab = 'Ozon (ppb)', main = 'Box-plot obelezja ozon u Njujorku')
> text(locator(1), "Outlier", adj=0)
```



**Box-plot obeležja ozon u Njujorku**



```
> set.seed(3)
> plot(runif(10), 1:10)
> locator(3, type="l")
$x
[1] 0.2958562 0.5765726 0.3839241
$y
[1] 7.975773 9.012613 2.993881
```

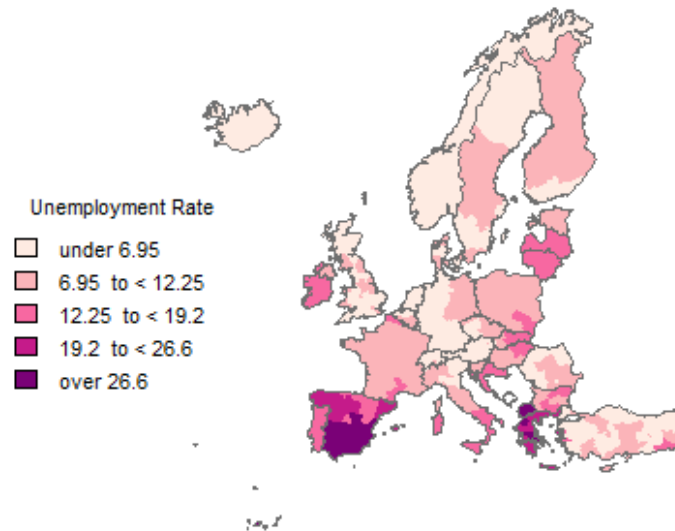
- `identify(x, y, labels)` - omogućeno je da korisnik istakne bilo koju tačku definisanu sa `x` i `y` (korišćenjem leve tipke miša) crtaњem odgovarajuće komponente iz `labels` (ili indeksnog broja tačke ako nisu navedene oznake) u blizini tačke.

```
> plot(Poverty ~ Unemp, data=usa.data)
> identify(x=usa.data$Unemp, y=usa.data$Poverty, labels=usa.data$State, n=3)
[1] 9 24 30
```





- Mapping data from Eurostat using R



## У презентацији коришћено:

- States03.csv - база података која садржи низ информација за сваку од 50 савезних држава у САД; подаци су прикупљени из више владиних извора и обједињени, а односе се на 2003. годину.
- подаци са сајта [epp.eurostat.ec.europa.eu/](http://epp.eurostat.ec.europa.eu/)  
Eurostat је статистичка канцеларија ЕУ, са седиштем у Луксембургу. Прикупља и објављује статистичке податке држава чланица ЕУ, држава кандидата, као и држава ЕФТА-е.