

# Korelacije

# Opšte karakteristike

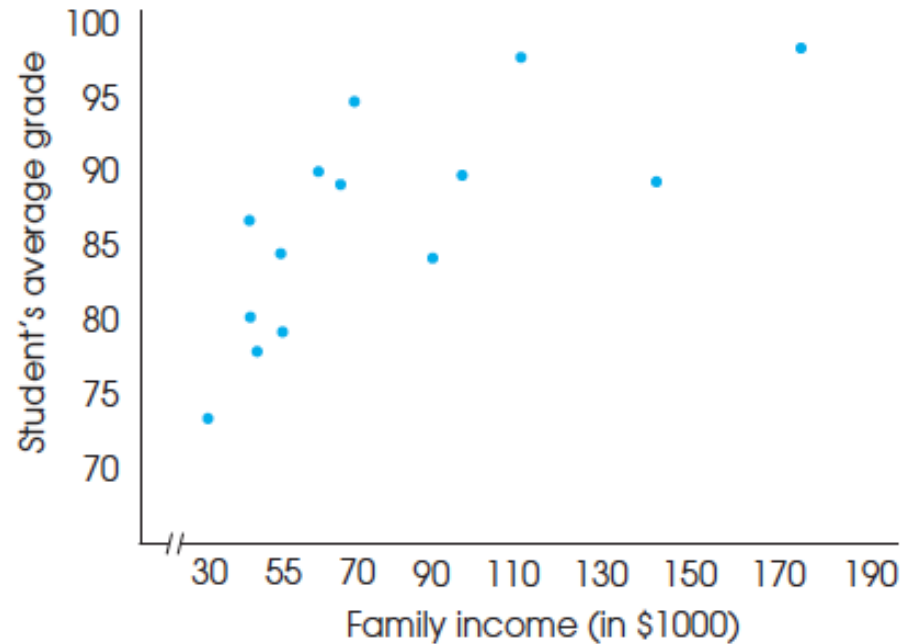
- ❑ Statistička tehnika koja se koristi da za opisivanje odnosa između **dvije varijable**
- ❑ Za računanje potrebna **dva seta vrijednosti** za svaku pojedinačnu opservaciju
- ❑ Intervalne mjere
- ❑ Korelacija je numerička vrijednost koja opisuje 3 karakteristike odnosa između dvije varijable (X i Y): ***smjer, oblik i snagu.***

# Smjer

- ❑ **SMJER:** obilježen je znakom ispred vrijednosti korelacije (pozitivna ili negativna)
- Pozitivna korelacija (+) – dvije pojave imaju tendenciju da se mijanjaju u istom smjeru. Kada raste X onda raste Y, ali sa smanjenjem X takođe vidimo i opadanje vrijednosti Y.
- Negativna korelacija (-) – dvije pojave imaju tendenciju da idu u suprotnom smjeru. Na primjer, sa porastom X smanjuje se Y.



Person	Family Income (in \$1000)	Student's Average Grade
A	31	72
B	38	86
C	42	81
D	44	78
E	49	85
F	56	80
G	58	91
H	65	89
I	70	94
J	90	83
K	92	90
L	106	97
M	135	89
N	174	95



**FIGURE 14.1**

Correlational data showing the relationship between family income ( $X$ ) and student grades ( $Y$ ) for a sample of  $n = 14$  high school students. The scores are listed in order from lowest to highest family income and are shown in a scatter plot.

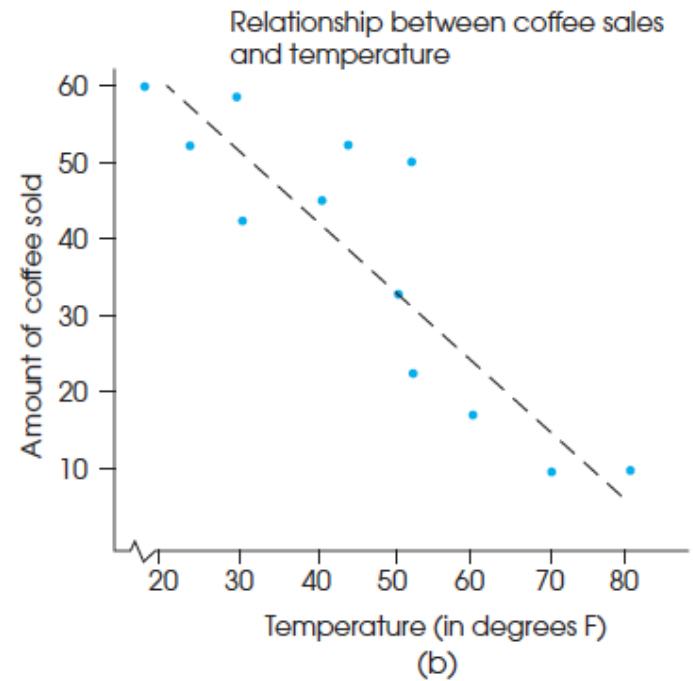
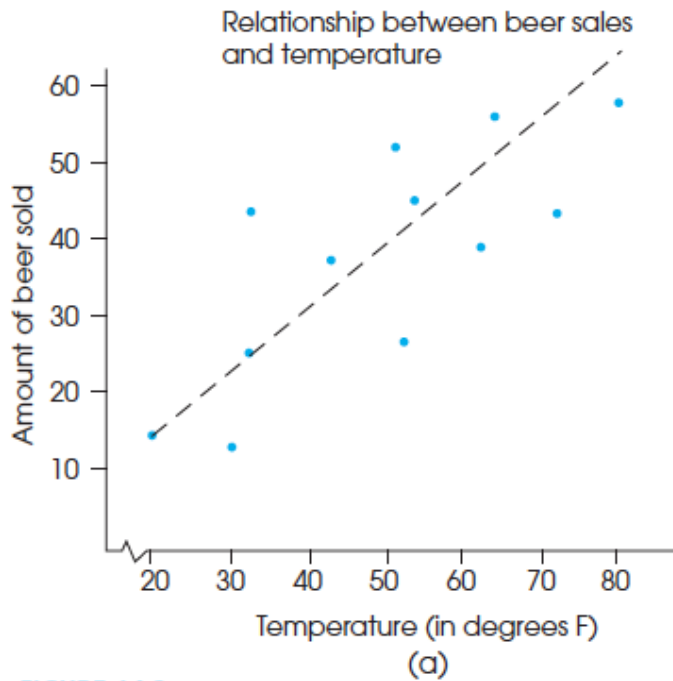
# Oblik

## □ OBLIK:

- Linearan odnos – vrijednosti imaju tendenciju da budu grupisane oko prave linije. Korelacije se najčešće koriste upravo za linearne relacije.
- Krivolinijski odnos – npr. vrijednosti se grupišu u obliku slova “U” ili obrnutog slova “U”. U slučaju ovakve distribucije, koeficijent korelacije je neprecizna mjera relacije između dvije varijable.



Univ



**FIGURE 14.2**

Examples of positive and negative relationships. (a) Beer sales are positively related to temperature. (b) Coffee sales are negatively related to temperature.

# Snaga odnosa

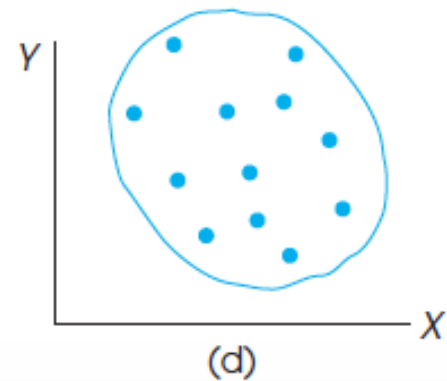
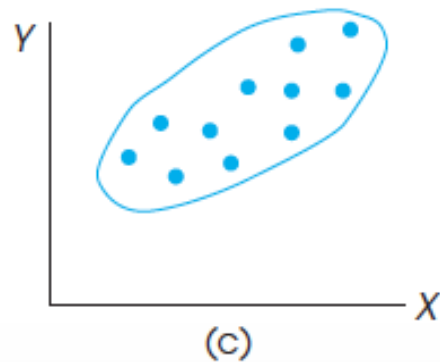
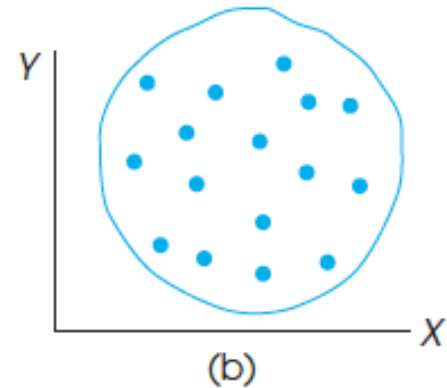
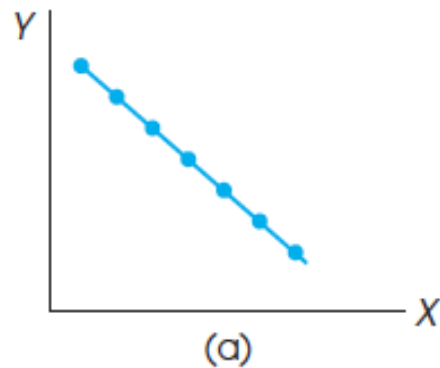
## □ **SNAGA:**

- Mjeri koliko je odnos između dvije varijable konsistentan. Drugim riječima, koliko prava linija precizno opisuje odnos
- Vrijednosti se kreću od  $-1$  do  $+1$
- Rijetko (nikad) smo u prilici da analiziramo savršenu korelaciju
- Nepovezanost između dvije pojave opisuje se vrijednošću  $0$



**FIGURE 14.3**

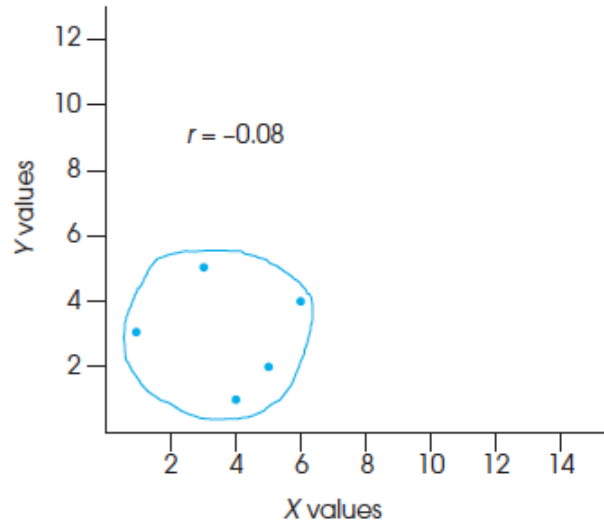
Examples of different values for linear correlations: (a) a perfect negative correlation,  $-1.00$ ; (b) no linear trend,  $0.00$ ; (c) a strong positive relationship, approximately  $+0.90$ ; (d) a relatively weak negative correlation, approximately  $-0.40$ .





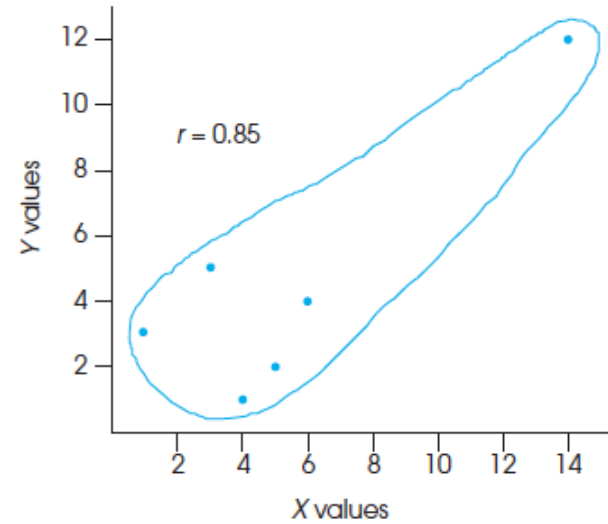


Univ



Original Data		
Subject	X	Y
A	1	3
B	3	5
C	6	4
D	4	1
E	5	2

(a)



Data with Outlier Included		
Subject	X	Y
A	1	3
B	3	5
C	6	4
D	4	1
E	5	2
F	14	12

(b)

FIGURE 14.7

A demonstration of how one extreme data point (an outlier) can influence the value of a correlation.

# Zadatak za vježbu

- Za sljedeću set vrijednosti nacrtaj tačkasti grafik i izračunaj Pirsonovu korelaciju.

<i>X</i>	<i>Y</i>
0	2
10	6
4	2
8	4
8	6

$$r = \frac{SS_{xy}}{\sqrt{SS_{xx}SS_{yy}}}$$

# Zadatak za vježbu

- Za sljedeću set vrijednosti nacrtaj tačkasti grafik i izračunaj Pirsonovu korelaciju.

<i>X</i>	<i>Y</i>
0	2
10	6
4	2
8	4
8	6

$$r = \frac{SS_{xy}}{\sqrt{SS_{xx}SS_{yy}}}$$

$$SS_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y})$$

# Zadatak za vježbu

- Za sljedeću set vrijednosti nacrtaj tačkasti grafik i izračunaj Pirsonovu korelaciju.

$X$	$Y$	$X - \bar{X}$	$Y - \bar{Y}$	$(X - \bar{X})^2$	$(Y - \bar{Y})^2$	$(X - \bar{X})(Y - \bar{Y})$
0	2	-6	-2	36	4	+12
10	6	+4	+2	16	4	+8
4	2	-2	-2	4	4	+4
8	4	+2	0	4	0	0
8	6	+2	+2	4	4	+4
				$SS_{XX} = 64$	$SS_{YY} = 16$	$SS_{XY} = 28$

# Zadatak za vježbu

- Za sljedeću set vrijednosti nacrtaj tačkasti grafik i izračunaj Pirsonovu korelaciju.

$X$	$Y$	$X - \bar{X}$	$Y - \bar{Y}$	$(X - \bar{X})^2$	$(Y - \bar{Y})^2$	$(X - \bar{X})(Y - \bar{Y})$
0	2	-6	-2	36	4	+12
10	6	+4	+2	16	4	+8
4	2	-2	-2	4	4	+4
8	4	+2	0	4	0	0
8	6	+2	+2	4	4	+4
				$SS_{XX} = 64$	$SS_{YY} = 16$	$SS_{XY} = 28$

$$R = 0.875$$