



Korelacije

METODOLOGIJA POLITIČKIH NAUKA

Opšte karakteristike

- ▶ Korelacije koristimo kada su i zavisna i nezavisna varijabla intervalne
- ▶ **Pirsonov koeficijent korelacije** – najčešće korišćena statistička tehnika u političkim naukama
- ▶ Označava se sa $-r$
- ▶ Za računanje potrebna **dva seta** vrijednosti za svaku pojedinačnu opservaciju
- ▶ Tri ključne karakteristike: **statistička zanačajnost, snaga odnosa i smjer odnosa**

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

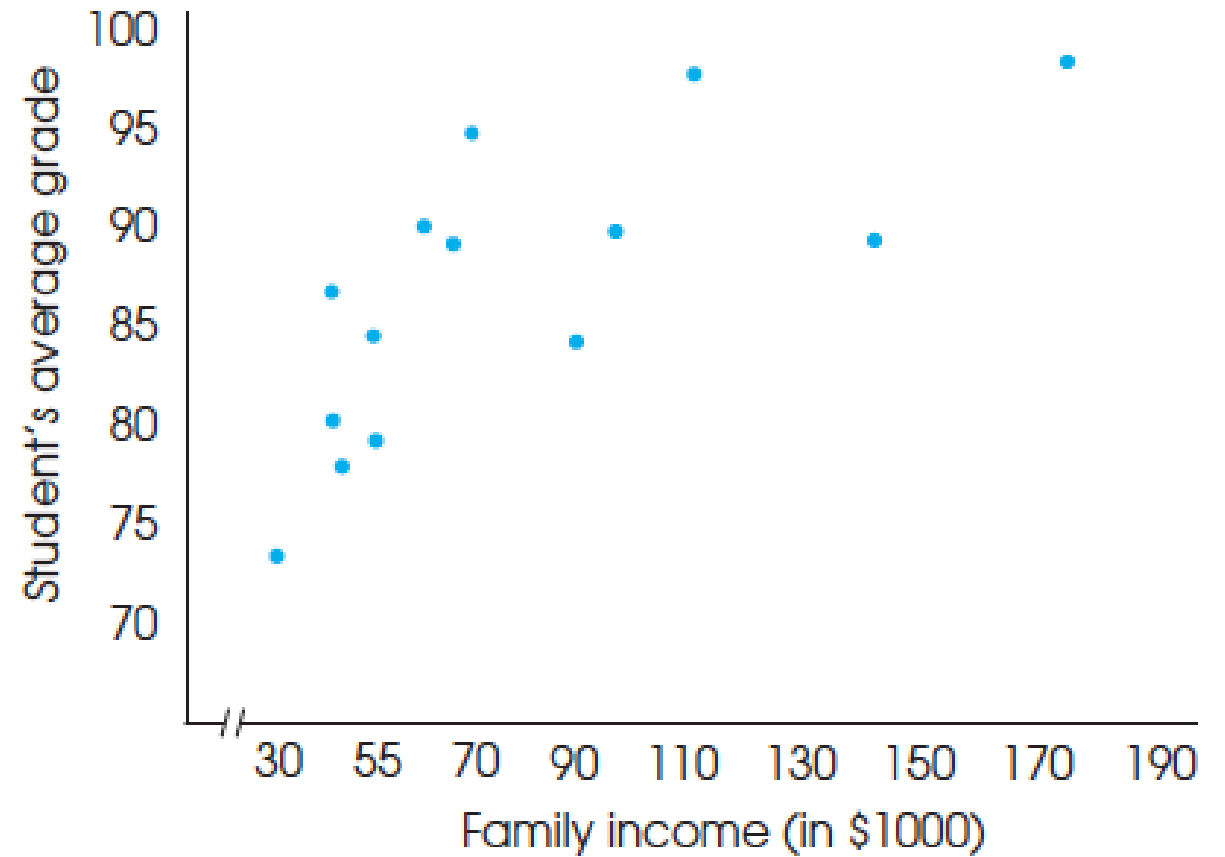


FIGURE 14.1

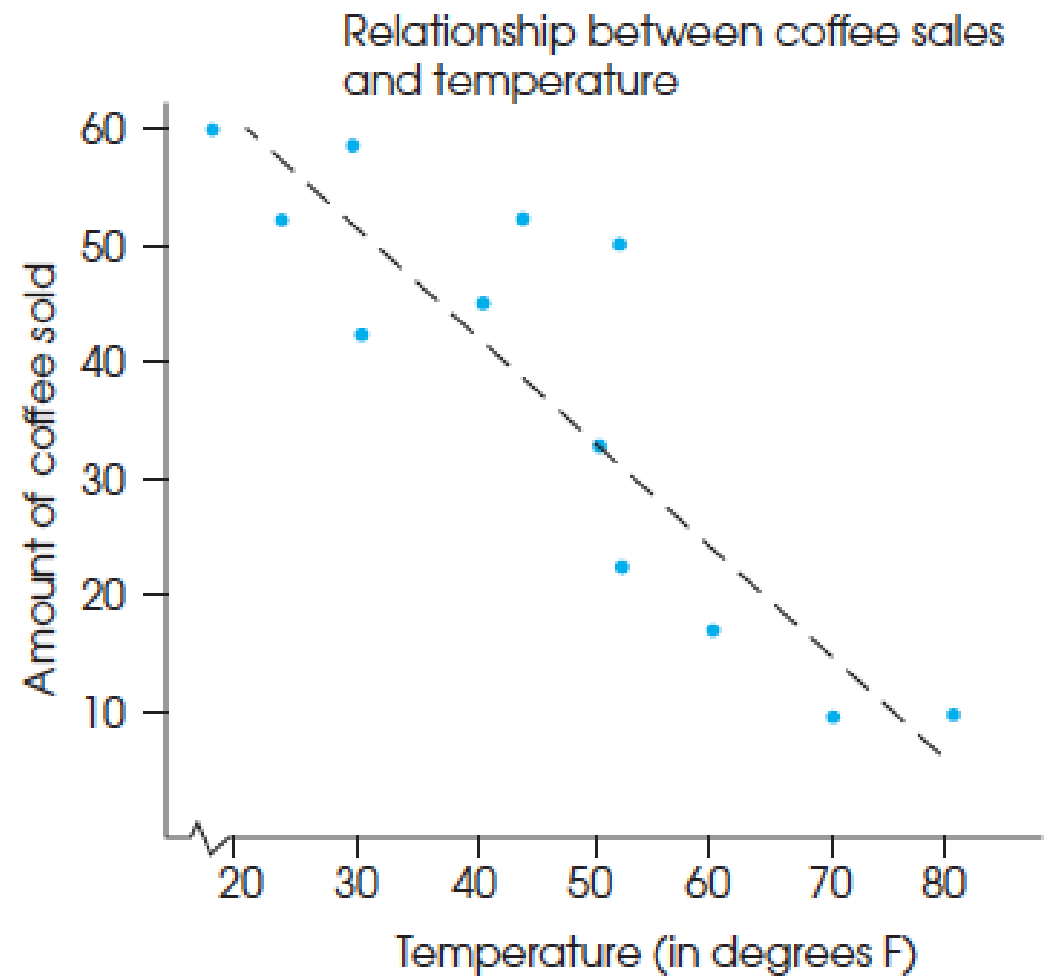
Correlational data showing the relationship between family income (X) and student grades (Y) for a sample of $n = 14$ high school students. The scores are listed in order from lowest to highest family income and are shown in a scatter plot.

Smjer odnosa

- ▶ Obilježen znakom ispred vrijednosti korelacije (pozitivna vs. negativna)
- ▶ **Pozitivna korelacija:** dvije pojave imaju tendenciju da variraju zajedno (u istom smjeru). Kada raste X onda Y, ali i sa smanjenjem X opada vrijednost Y.
- ▶ **Negativna korelacija:** dvije pojave imaju tendenciju da variraju u suprotnom smjeru. Sa porastom X smanjuje se Y.



(a)



(b)

FIGURE 14.2

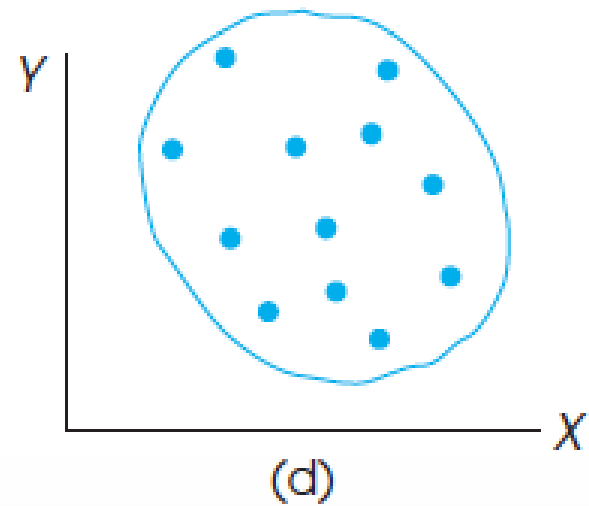
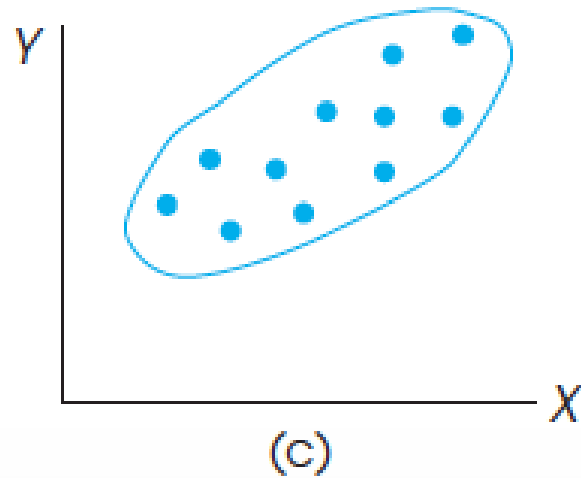
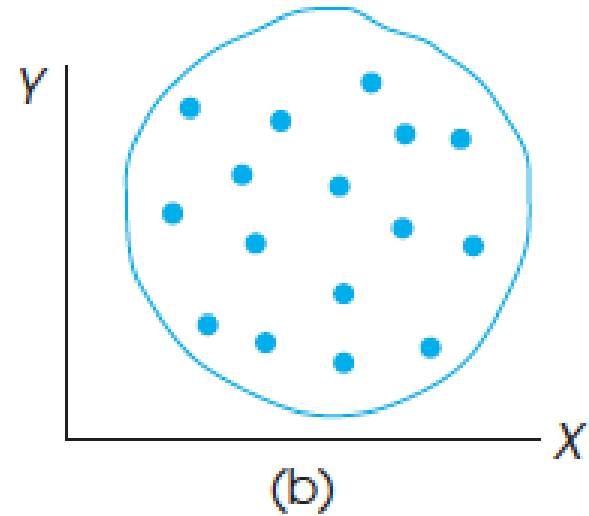
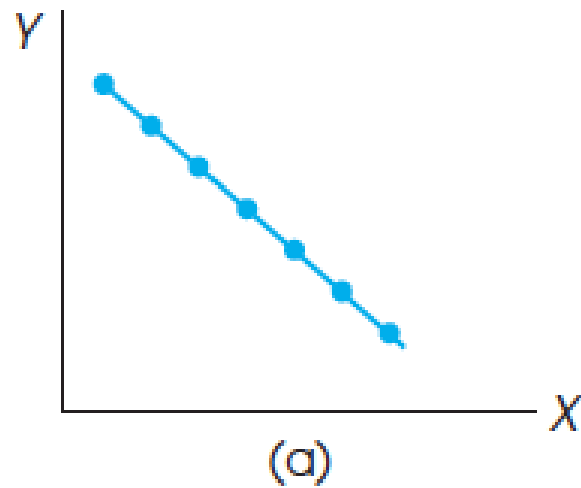
Examples of positive and negative relationships. (a) Beer sales are positively related to temperature. (b) Coffee sales are negatively related to temperature.

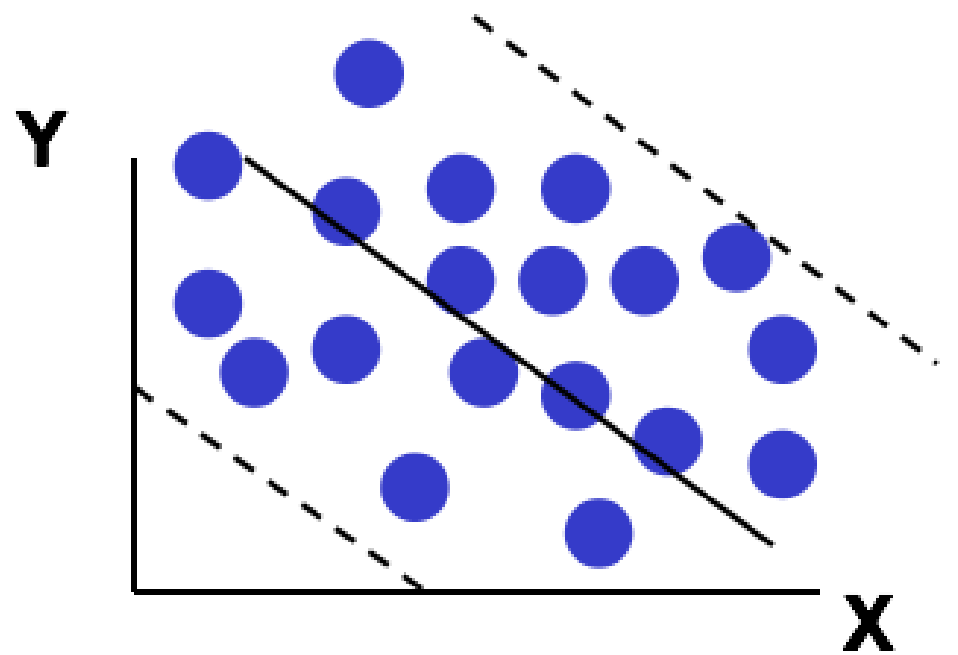
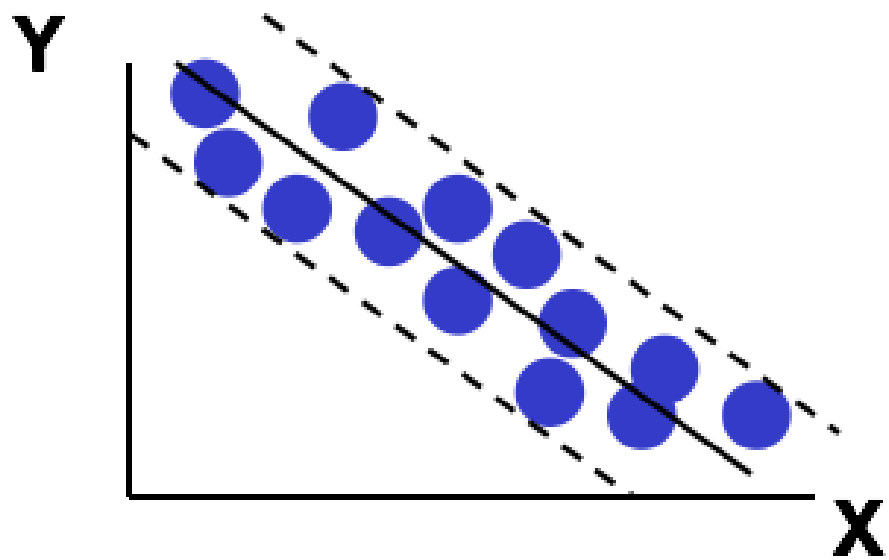
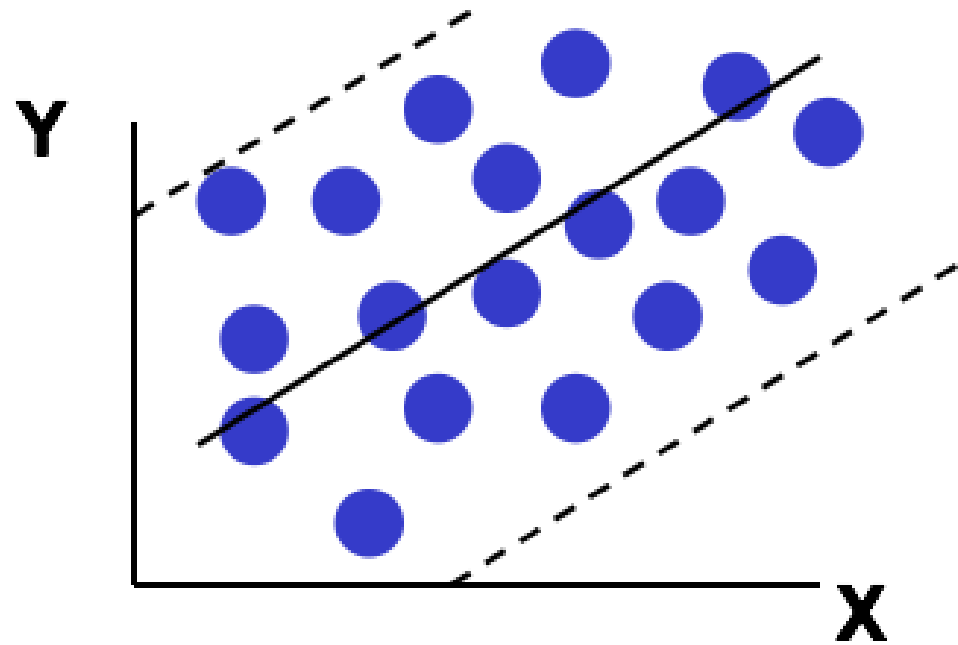
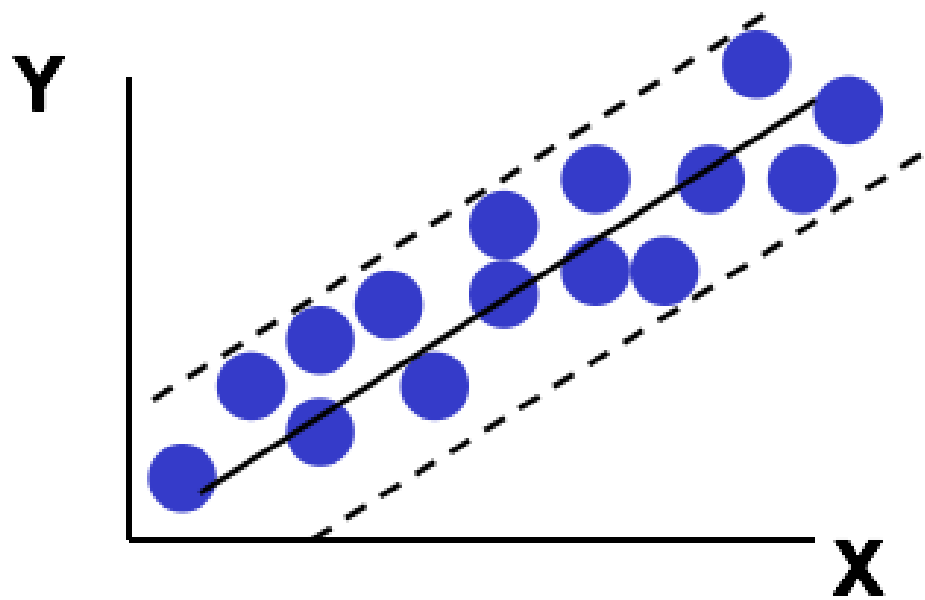
Snaga odnosa

- ▶ Koliko prava linija dobro opisuje odnos između dvije pojave
- ▶ Vrijednost korelacije se kreću **od -1 do +1**
- ▶ **Nepostojanje veze** između dvije pojave se označava 0
- ▶ Rijetko (nikad) nijesmo u prilici da analiziramo savršenu korelaciju
- ▶ Ne postoji ustaljeno pravilo za **ocjenu snage odnosa**, ali najčešće:
 - Slaba: od 0 - 0.25
 - Umjerena: od 0.25 do 0.60
 - Snažna: od 0.60 do 1
- ▶ Što je **manji uzorak** to je veći koeficijent potreban da se dostigne statistička značajnost

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 .

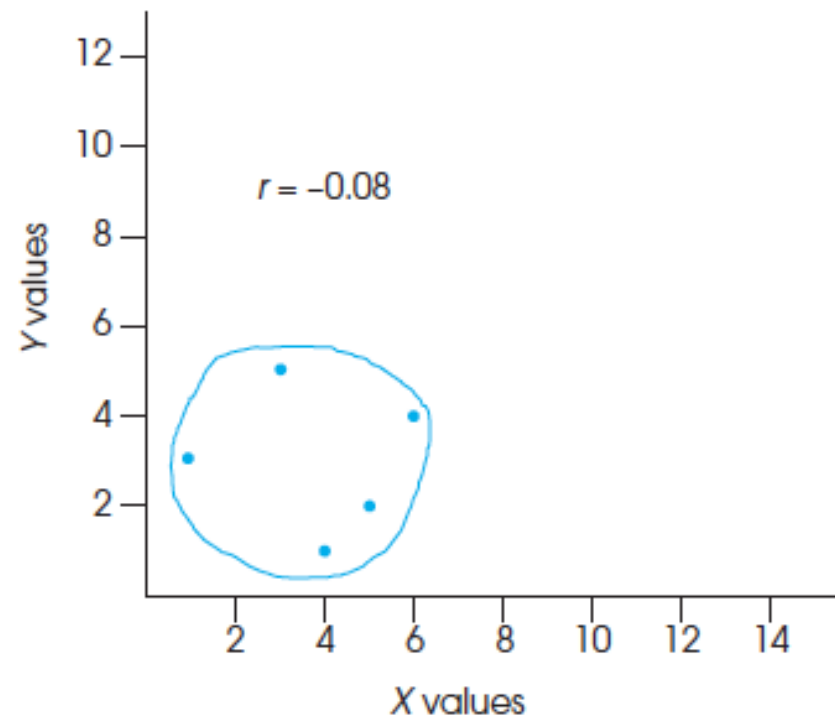




Interpretiranje ključnih informacija iz korelacionih matrica

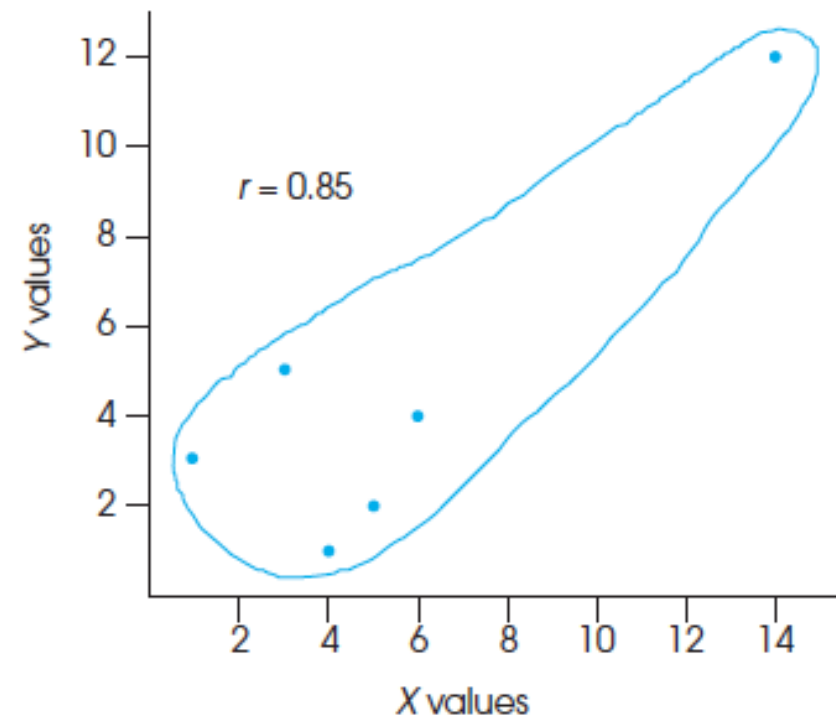
Tabela 7.2 Korelacijska matrica skorova na termometru osjećanja

	Gor	Buš	Feministkinje	Armija
Al Gor				
<i>r</i>	1	-.414**	.331**	-.072**
Stat. znač.		.000	.000	.005
<i>N</i>	1774	1747	1410	1494
Džordž V. Buš				
<i>r</i>	-.414**	1	-.203**	.270**
Stat. znač.	.000		.000	.000
<i>N</i>	1774	1761	1403	1487
Feministkinje				
<i>r</i>	.331**	-.203**	1	.033
Stat. znač.	.000	.000		.219
<i>N</i>	1410	1403	1427	1413
Armija				
<i>r</i>	-.072**	.270**	.033	1
Stat. znač.	.005	.000	.219	
<i>N</i>	1494	1487	1413	1517



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

(a)

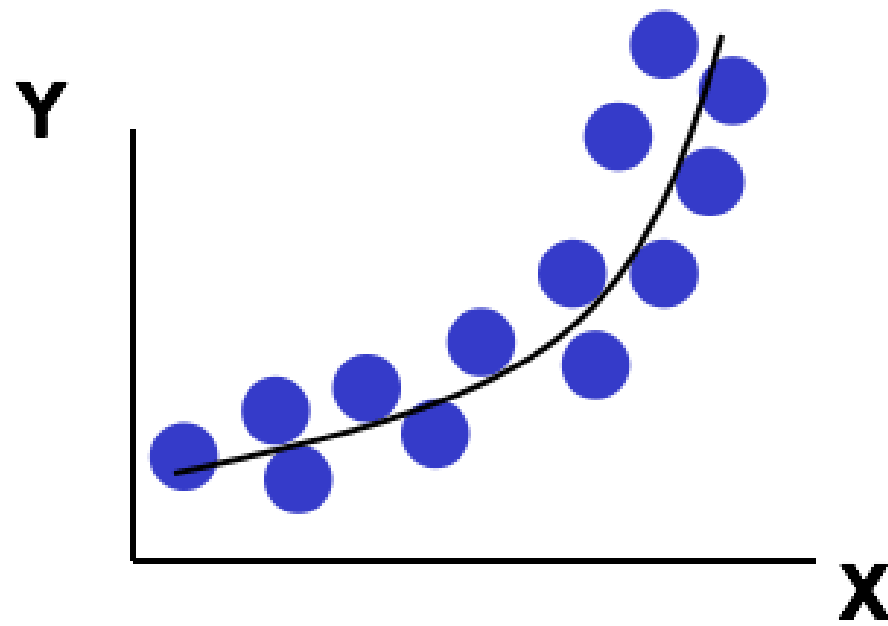
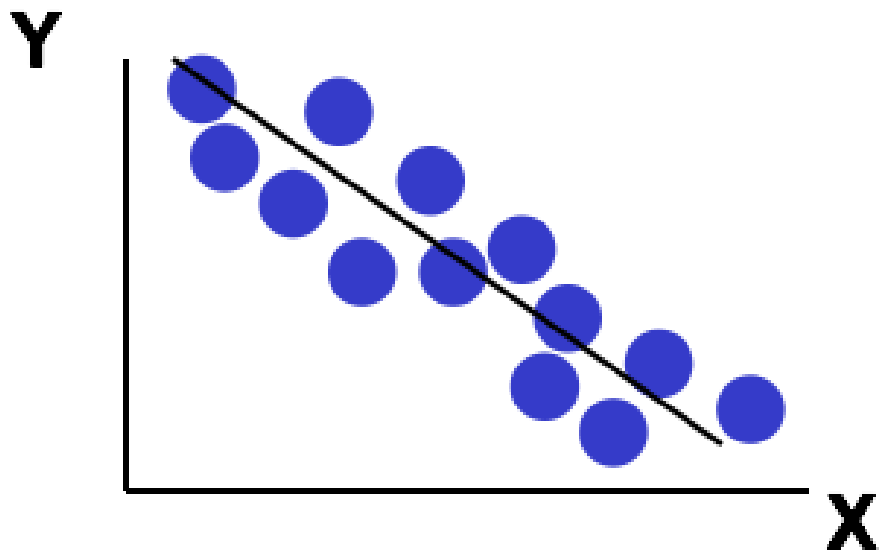
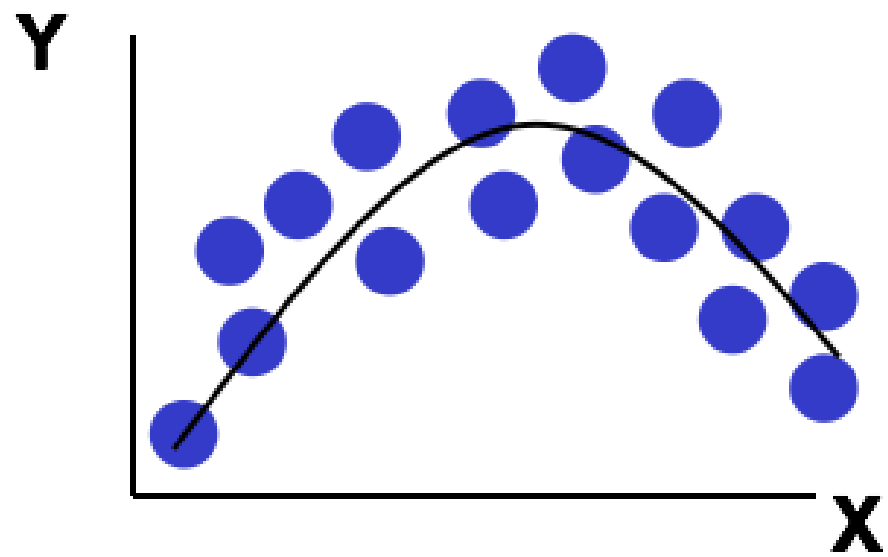
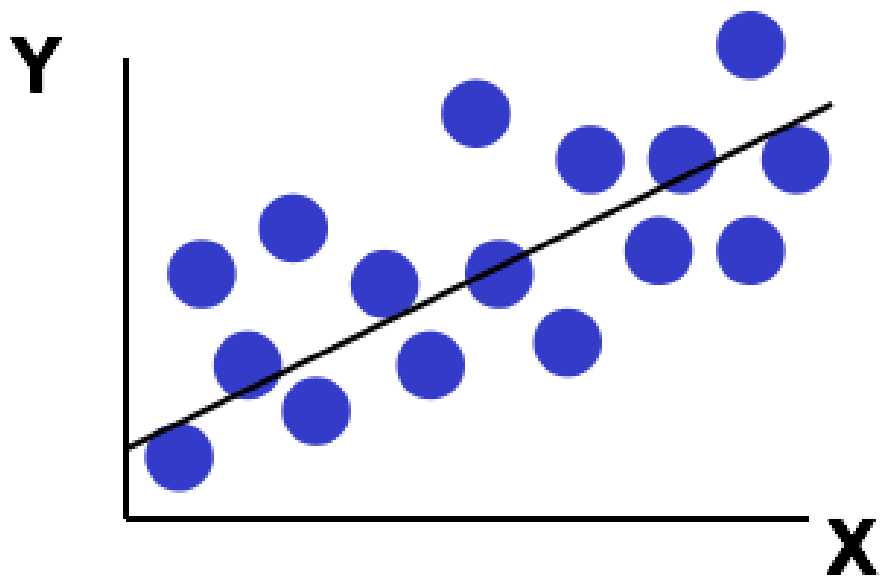


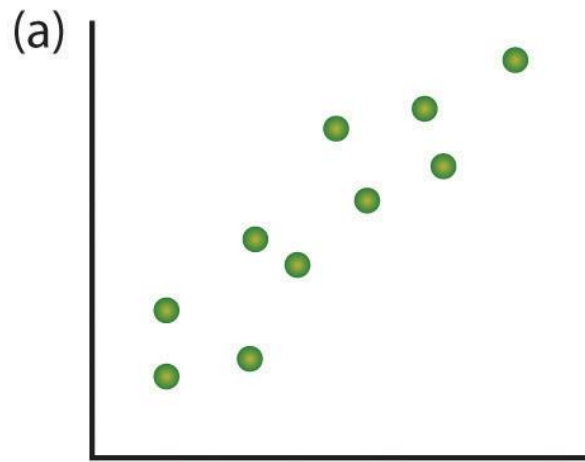
Data with Outlier Included		
Subject	X	Y
A	1	3
B	3	5
C	6	4
D	4	1
E	5	2
F	14	12

(b)

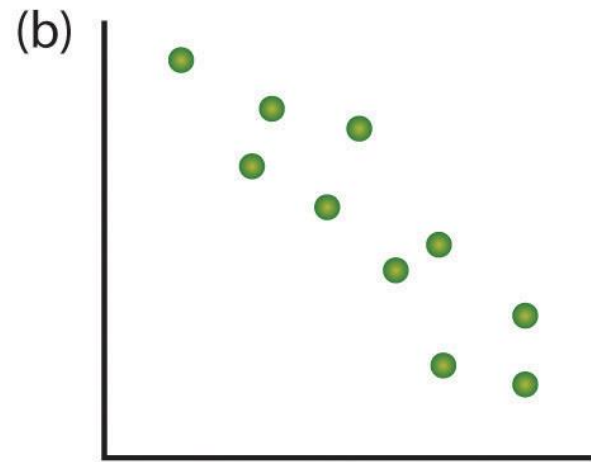
FIGURE 14.7

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

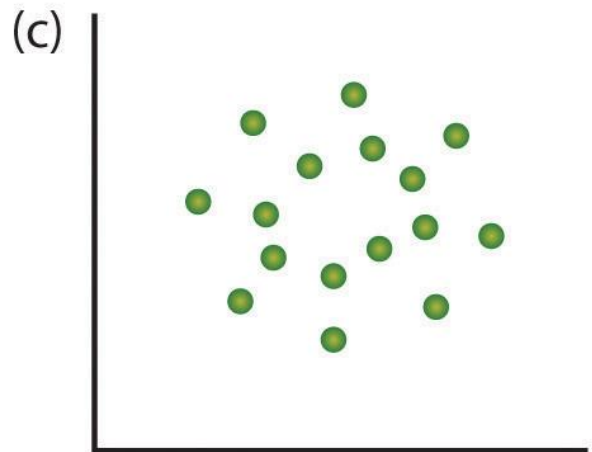




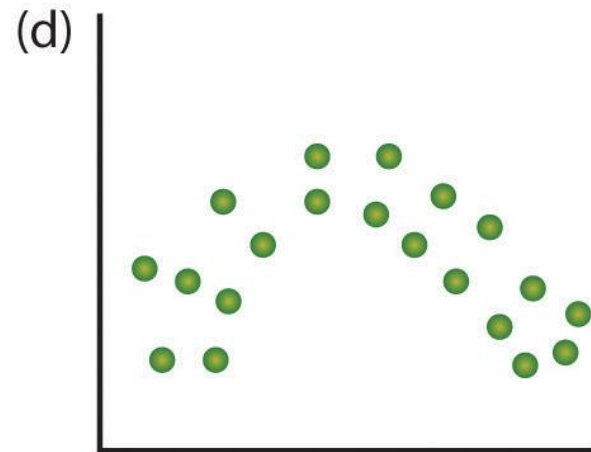
Positive linear
 $r = +.82$



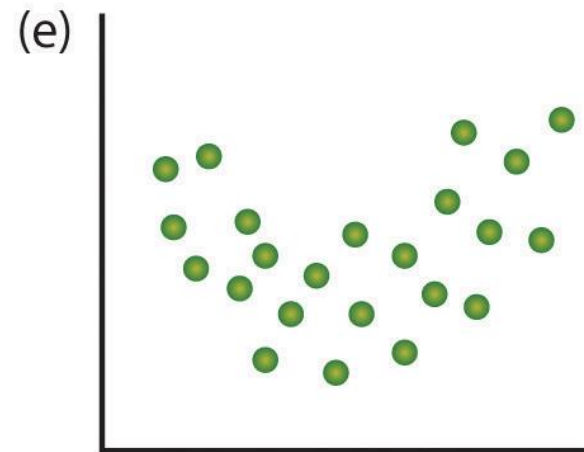
Negative linear
 $r = -.70$



Independent
 $r = 0.00$



Curvilinear
 $r = 0.00$



Curvilinear
 $r = 0.00$



Linearna regresija

METODOLOGIJA POLITIČKIH NAUKA

Opšte karakteristike

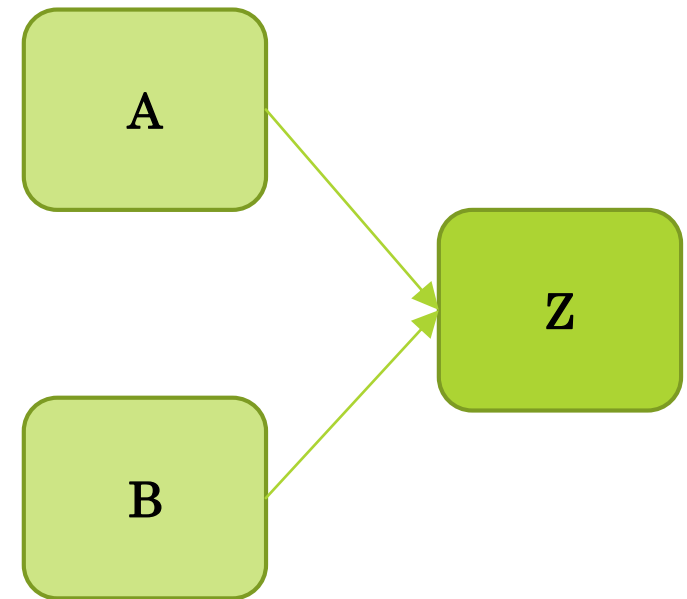
- ▶ Korelacija govori samo (i isključivo) o postojanju odnosa između dva fenomena
- ▶ **Prosta linearna regresija (PLR):** prvi korak ka **kauzalnosti** (i dalje veoma daleko...)
- ▶ Zahtijeva od nas da specifikujemo koja je varijabla nezavisna a koja zavisna
- ▶ **Višestruka linearna regresija (VLR):** drugi korak ka kauzalnosti
- ▶ Regresija je ubjedljivo najčešće korišćeni metod u društvenim naukama

Višestruka linearna regresija

- ▶ Prava vrijednost statističkih analiza dolazi iz mogućnosti da analiziramo tri ili više faktora istovremeno
- ▶ “Rješenje” za problem lažne povezanosti – vrijednost **kontrola**
- ▶ **Kontrolne varijable** – nezavisne varijable, ali suštinski (teorijski) manje važne
- ▶ Kontrolisanje kod krostabulacija? - najčešće najviše tri varijable
- ▶ Komparativna ocjena uticaja raznih faktora

Korisne informacije iz VLR

- ▶ VLR će procijeniti uticaj varijable A na Z istovremeno sa uticajem B na Z
- ▶ Koeficijent efekta A - Z će biti procijenjen dok kontolišemo za B (uslobljen efekat)
- ▶ Za oba efekta možemo dobiti informaciju o statističkoj značajnosti, smjeru u snazi odnosa
- ▶ Model nam može reći i koliki procenat varijacije na nezavisnoj varijabli možemo objasniti nezavisnim varijablama – **R kvadrat**



R kvadrat

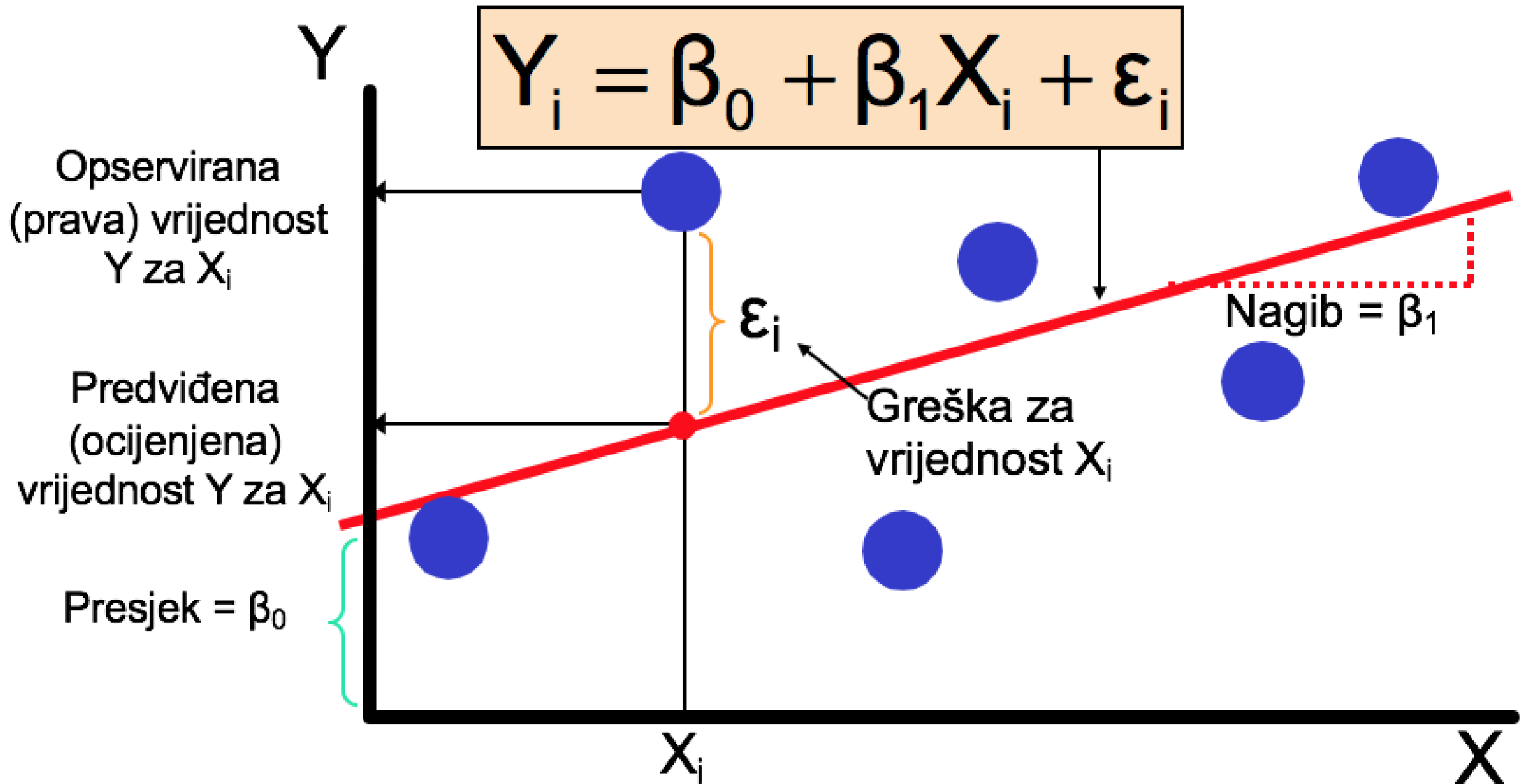
- ▶ Ocjena **ukupne snage** modela
- ▶ Može uzeti vrijednosti **od 0 do 1**
- ▶ Interpretira se u **procentima**
- ▶ R kvadrat **uvijek** raste sa porastom broja varijabli
- ▶ Zadovoljavajuća vrijednost u političkim naukama varira od poddiscipline

Model regresije (Formula)

The diagram illustrates the regression formula $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ with the following labels and annotations:

- Zavisna varijabla**: Points to Y_i .
- Presjek Y**: Points to β_0 .
- Koeficijent nagiba**: Points to β_1 .
- Nezavisna varijabla**: Points to X_i .
- Slučajna greška**: Points to ε_i .
- Linearna komponenta**: A blue bracket under $\beta_0 + \beta_1 X_i$.
- Komponenta greške**: A blue bracket under ε_i .

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$



Interpretacija

- ▶ **Presjek:** b_0 je ocijenjena prosječna vrijednost Y kad je X jednaka nuli
 - Primjer?
 - Nema uvijek intuitivnog smisla
- ▶ **Nagib:** b_1 je ocijenjena promjena prosječne vrijednosti Y kao rezultat jedinične promjene X
 - Standardizovani koeficijenti? – std. devijacija postaje jedinica mjere

Važna pitanja

- ▶ Koliko varijabli uključiti u model?
- ▶ Šta ako je teorijski značajna varijabla statistički beznačajna?
- ▶ Da li je zavisna varijabla suštinski različita od nezavisne varijable?
- ▶ Interakcije?
- ▶ Ostali tipovi regresija – logistička regresija (*dummy variable*)

Pretpostavke

- ▶ Intervalni podaci (makar 5-kategorija)
- ▶ Sveobuhvatan model (*Full model specification*)
- ▶ Normalna distribucija
- ▶ Linearan odnos
- ▶ Nezavisne opservacije

Tabela 7.3. Višestruka regresija: Stav prema klimatskim promjenama

	Vjerovanje u klimatske promjene	Klimatske promjene uzrokovane ljudskim faktorom	Neophodnost djelovanja	Obnovljiva energija
Ideologija	-.483**	-.671**	-.542**	-.305*
Individualizam	-.051	-.128**	-.116**	.067
Hijerarhija	.103*	.176**	.111**	.095
Egalitarizam	.092*	.171**	.244**	.021
Fatalizam	-.032	-.079	-.150**	.009
Obrazovanje	-.025	.025	.055	-.137
Pol	-.331	-.369	-.355	-.814*
Svijest o klimatskim promjenama	.339'	.239'	.308*	-.009
Pril. R ²	.180	.388	.395	.061
N	262	262	263	255

Izvor: Michael D. Jones, "Leading the Way to Compromise? Cultural Theory and Climate Change Opinion," *PS: Political Science and Politics* 44, no. 4 (October 2011): 720–25.

Napomena: Koeficijenti nisu standardizovani

*Značajna na nivou .05

** Značajna na nivou .01

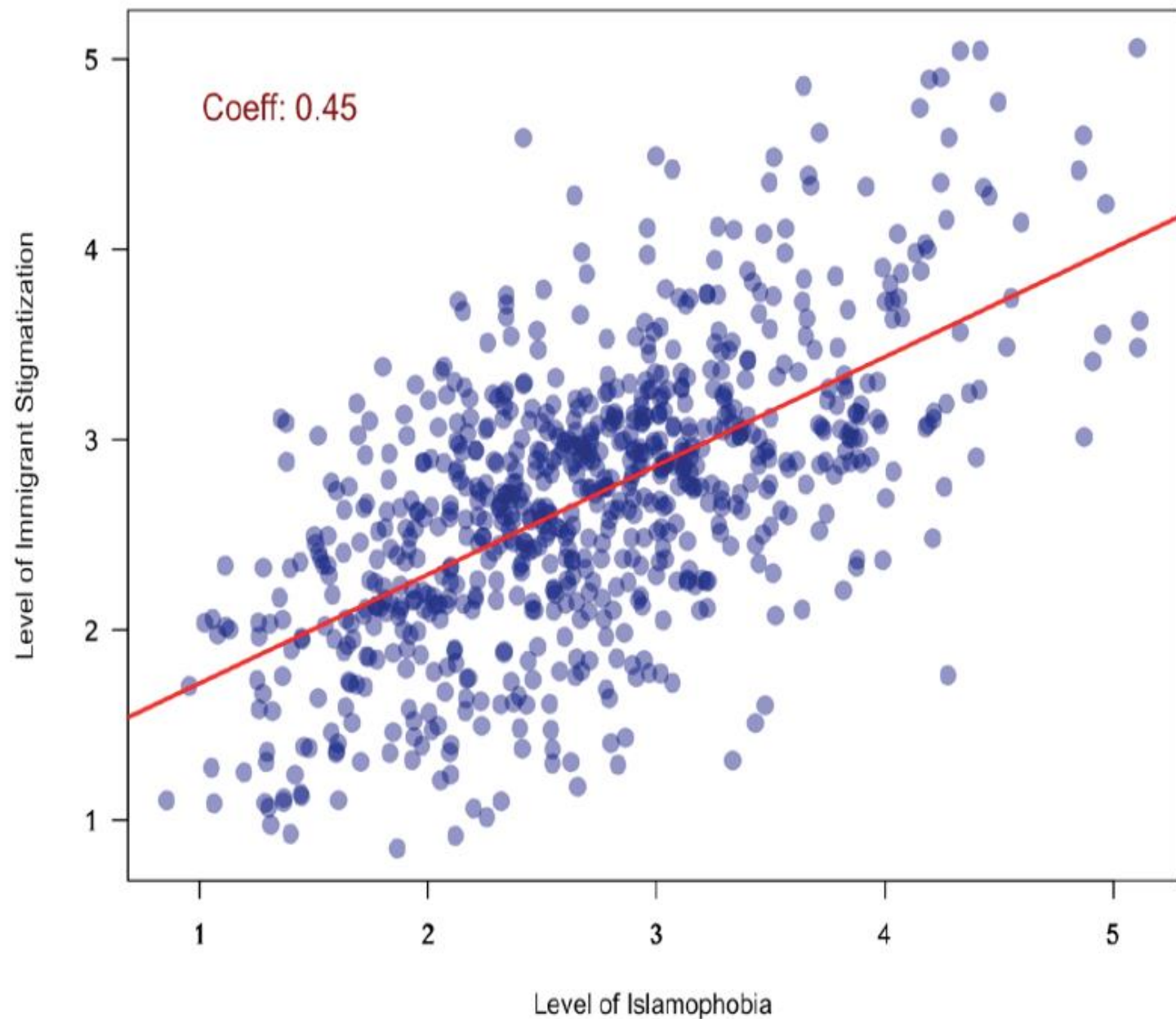
Table 1: Ordinary Least Squares (OLS) Regression Results

	<i>DV: Islamophobia</i>		
	Full Sample	EU	WB
	(1)	(2)	(3)
<i>Intercept</i>	1.514*** (0.270)	1.789*** (0.401)	1.462*** (0.356)
<i>Education</i>	-0.052** (0.026)	-0.147*** (0.039)	0.021 (0.034)
<i>Islam Knowledge</i>	-0.043* (0.021)	-0.037 (0.032)	-0.063** (0.027)
<i>National Identity</i>	0.101*** (0.028)	0.119*** (0.042)	0.082** (0.036)
<i>Religiousness</i>	0.023 (0.022)	0.063** (0.029)	-0.012 (0.032)
<i>Left-Right</i>	0.086*** (0.017)	0.104*** (0.026)	0.055** (0.021)
N	568	243	325
R ²	0.217	0.339	0.186
Adj. R ²	0.197	0.298	0.149
Res. SE	0.712 (df = 553)	0.685 (df = 228)	0.678 (df = 310)
F Stat.	10.928***	8.344***	5.061***

Note:

*p<0.1; **p<0.05; ***p<0.01

The Effect of Islamophobia on Immigrant Stigma



Grafički prikaz uticaja islamofobije na nivo stigmatizacije imigranata



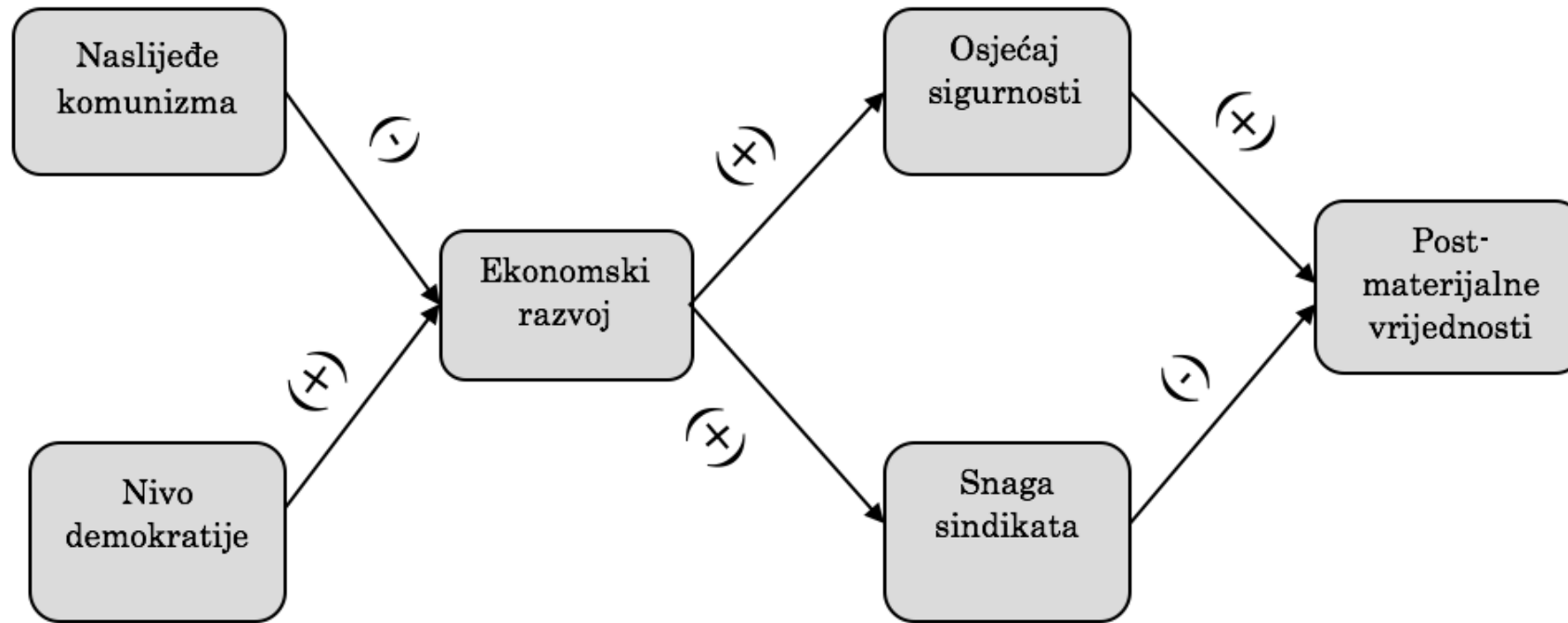
Domaći rad II

METODOLOGIJA POLITIČKIH NAUKA

1. Operacionalizuj pojam – *političko znanje*. Ponuditi definiciju koja je po vašem mišljenju odgovarajuća, odrediti najmanje dvije dimenzije ovog koncepta i za svaku od njih ponuditi po dva indikatora. Za svaki od indikatora navesti tip skale mjerenja (nominalna, ordinalna, intervalna) koja bi bila korišćena za mjerenje.

KONCEPT	POLITIČKO ZNANJE	
DEFINICIJA		
DIMENZIJE	Dimenzija 1:	Dimenzija 2:
INDIKATOR 1		
INDIKATOR 2		
SKALA (IND. 1)		

2. „Prevesti“ dijagram u koherentnu teoriju. Obratiti pažnju na ulogu (poziciju) varijable u modelu i interpretirati smjer odnosa sa varijablama koje je okružuju.



Unijeti odgovor:

3. Za svaki par navedenih varijabli ponuditi mogući kauzalni mehanizam sa jednom (ili dvije) posredujuće varijable. Objasniti teoriju na kojoj je baziran Vaš odgovor - *na koji način pomenuta varijabla/e posreduje/u između nezavisne i zavisne varijable?*

NEZAVISNA VARIJABLA	POSREDUJUĆA/E VARIJABLA/E	ZAVISNA VARIJABLA
<i>Nuklearna katastrofa</i>	(Navesti varijablu i objašnjenje)	<i>Glasanje za „zelene“ partije</i>
<i>Dužina radne nedelje</i>		<i>Kvalitet života u zajednici</i>
<i>Prirodni resursi</i>		<i>Autokratski režim</i>

4. Studije o ponašanju biračkog tijela prilično su uobičajene u političkim naukama, a odziv glasača je vjerovatno najistaknutija zavisna varijabla u ovoj literaturi. Pretpostavimo da istražujete izlaznost birača na kontinentu Lemurija, koji se sastoji od dvadeset i pet zemalja, ali je rijetko proučavan. Iz vašeg pregleda literature o glasanju u drugim dijelovima svijeta, saznajete da viši nivo prihoda i obrazovanja obično znači i veću izlaznost. Pronađite precizan odnos mjera ovih varijabli za 2010. godinu. Stroži uslovi registraciju, s druge strane, obično znače manju izlaznost.

	<i>Prihod glavi stanovnika</i>	<i>% stanovništva po sa srednjom stručnom spremom</i>	<i>Preduslovi za registraciju</i>	<i>Izlaznost u %</i>
Awkland	42.669	69	Srednji	90
Bluebell	14.047	54	Veoma visoki	56
Bonquerres	49.754	84	Srednji	84
Chinet	27.160	40	Visoki	63
Coliistan	15.889	70	Niski	63
Corazonia	5.626	32	Niski	80
Drago	12.455	34	Srednji	60
Flamboya	44.090	31	Srednji	68
Fluoristan	29.709	79	Niski	85

5. Interpretiraj numeričke tabele i grafike iz postojećih publikacija. Fokusirati se na tri ključne informacije: *statističku značajnost, snagu odnosa, i smjer odnosa* među varijablama.

	DPS	SDP	SNP	NOVA	PZP	Apstinenti	Ostali	
Muški	233	22	105	27	48	100	63	598
	39.0%	3.7%	17.6%	4.5%	8.0%	16.7%	10.5%	100.0%
Ženski	238	30	106	24	29	151	52	630
	37.8%	4.8%	16.8%	3.8%	4.6%	24.0%	8.3%	100.0%
Ukupno	471	52	211	51	77	251	115	1228
	38.4%	4.2%	17.2%	4.2%	6.3%	20.4%	9.4%	100.0%

Napomena: $X^2(6, N = 1228) = 16.746; p < .01$

6. Istraživač, Lizbet Hohe, je željela da utvrdi da li je rad za međunarodnu organizaciju uticao na podršku jedne osobe određenim nadnacionalnim normama. Ako jeste, da li bi se ta veza održala, kontrolišući druge moguće uticaje? Ona je anketirala najviše zvaničnike Evropske komisije o njihovim stavovima prema Evropskoj uniji, a matrica korelacije ispod pokazuje samo mali dio ovih rezultata. Kako biste tumačili ove bivarijatne odnose? Razmišljajte u kontekstu statističkog značaja, pravca i snage.

	Nadnacionalizam	Ideologija	Veličina zemlje	Međunarodno obrazovanje	Radni vijek
Nadnacionalizam	1	-.17*	.01	.22**	.16*
Ideologija		1	-.14	-.05	-.03
Veličina zemlje			1	-.04	.40**
Int. obrazovanje				1	.01
Radni vijek					1

*p < .05.

**p < .01.

7. Kako biste tumačili sljedeće rezultate iz modela višestruke regresije u kojoj je zavisna varijabla samoprijavljena politička ideologija, od ljevičarske (1) do desničarske (10), u Brazilu?

Koeficijenti	Nestandard. <i>B</i>	Standardizovano (Beta)	Znač.
Godine	.016	.091	.003
Dohodak	-.035	-.026	.393
Nivo obrazovanja	-.128	-.118	.000
Ukupan model	Kvadrat R=.032		

Izvor: World Values Survey wave 6 (data for Brazil, 2014).

Unijeti odgovor: