

KVANTITATIVNE METODE I

**Kineziološki fakultet
Sveučilište u Splitu
2008/2009**

Modul 2

Uvod u programski sustav Statistica 7.0

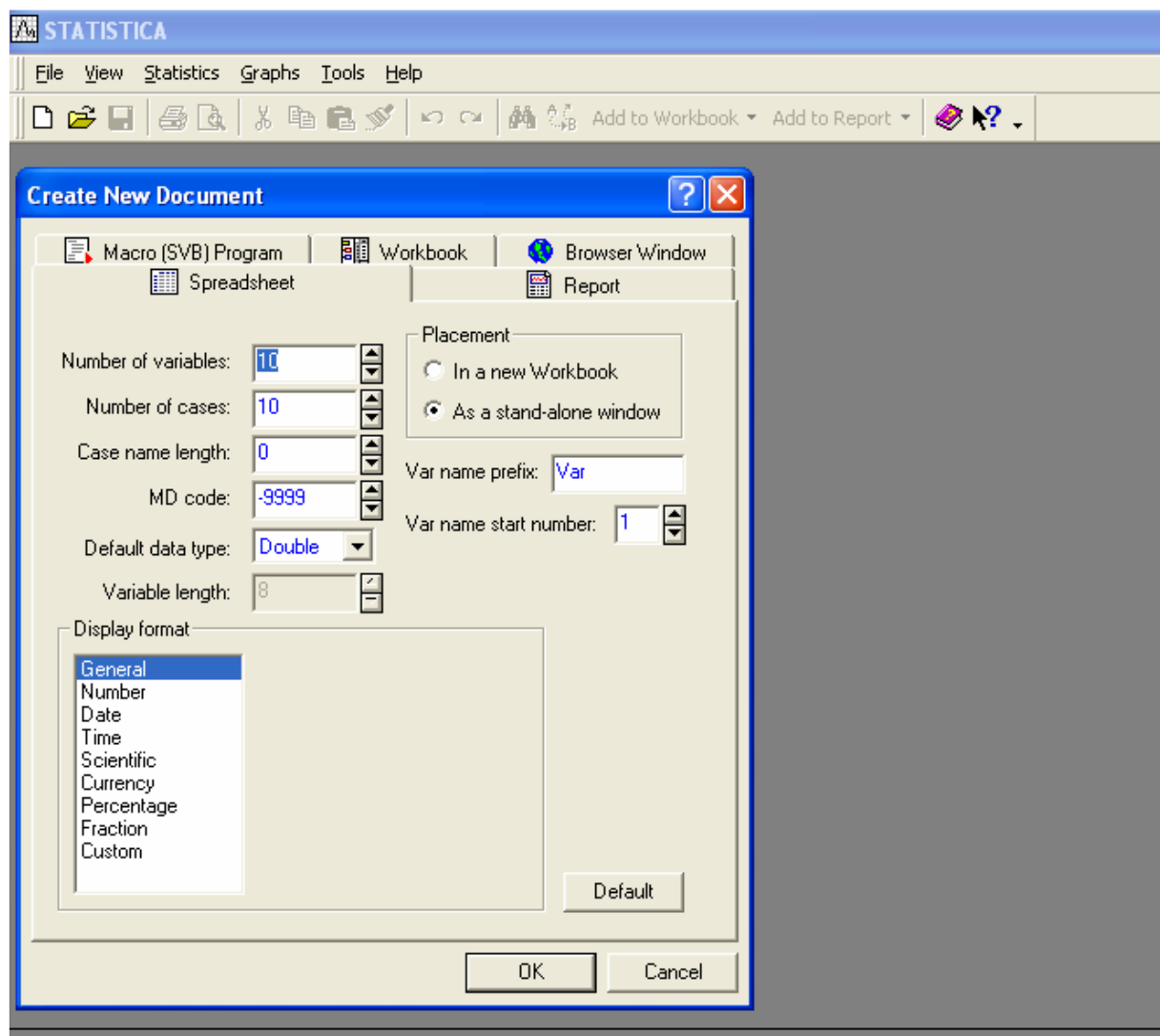
(radni materijali)

Priredio : mr.sc. Jelaska Igor

Uvod

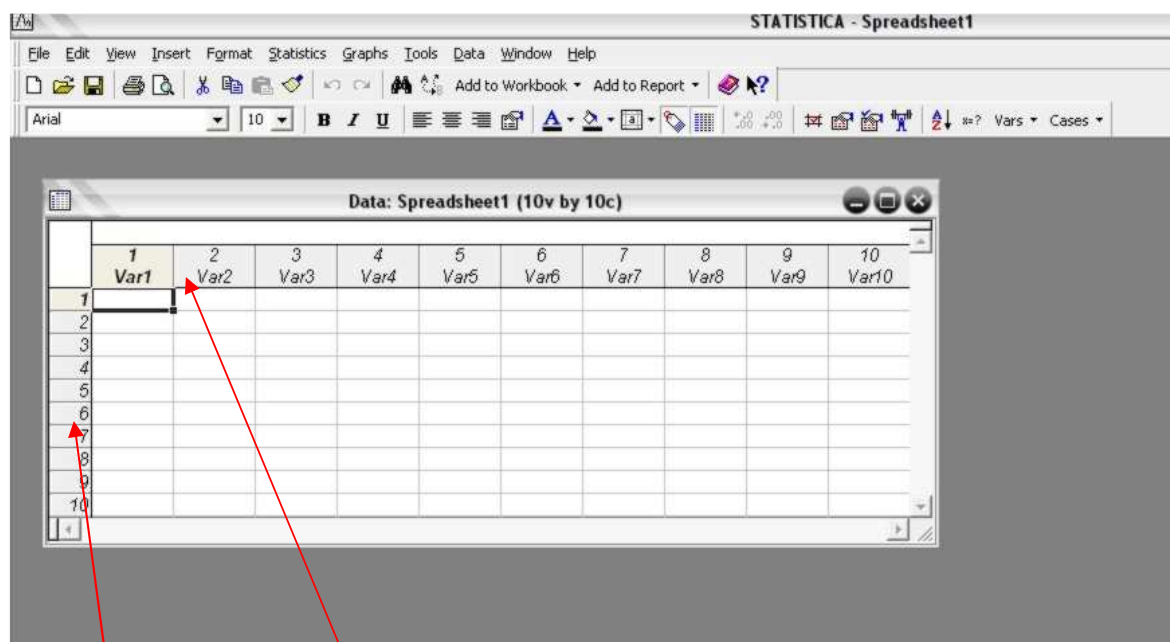
U današnje vrijeme statistička analiza je nezamisliva bez elektroničkog računala i prikladnog softvera. Kao što znamo, kineziologija kao znanost neke od zaključaka donosi na temelju rezultata statističkih proračuna. Stoga svaki diplomirani kineziolog mora savladati praktični i teoretski aspekt statistike, tj. treba znati statistički obraditi podatke i interpretirati dobivene rezultate.

Pokretanjem Statistice 7.0 (Start->All Programs->Statistica 7)javlja nam se slijedeće sučelje:



Number of variables – Broj varijabli koje istražujemo (stupci)
 Number of cases – broj entiteta
 Case name length – broj znakova za dužinu imena entiteta
 MD Code - missing data code, kod za vrijednosti koje nedostaju
 Default data type : tip podataka
 Variable length : duljina varijable, broj znakova varijable
 Var name prefix : defaultni prefix imena svake varijable
 Var name start number : numeracija varijabli polazi od tog broja

Sada nam se javlja:



ISPITANICI (CASES) **VARIJABLE (VARIBLES)**
 - U redcima - U stupcima

Upisivanje imena varijabla:

- 2 klika lijevom tipkom miša na vrh stupca

Upisivanje imena ispitanicima:

- 2 klika lijevom tipkom miša na početak retka

Upisivanje brojeva:

- decimalna mjesta se odvajaju zarezom; npr. 2,98 a ~~ne točkom~~ 2.98

Povećanje i smanjenje broja decimalnih mjesta:



- označiti stupac na koji želimo primijeniti određenu radnju, te pritisnuti
 - o za smanjenje decimalnih mjesta
 - o za povećanje decimalnih mjesta

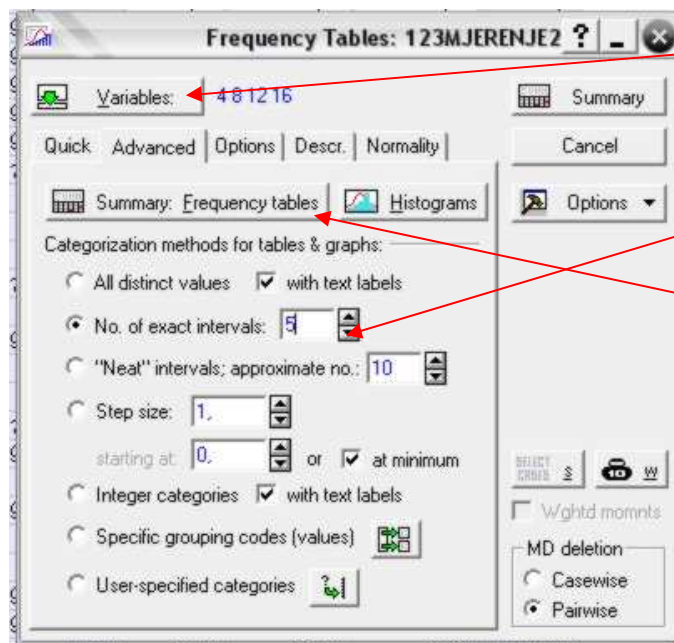
	1 SDM	2 BML	3 KUS								
1	148,33	6,30	9,45								
2	170,00	4,83	10,21								
3	170,67	5,72	9,85								
4	170,00	5,13	9,30								
5	161,67	5,67	9,51								
6	173,33	6,22	9,45								
7	167,67	6,07	9,63								
8	150,00	5,70	9,71								
9	159,33	6,58	8,82								
10	163,33	4,47	9,04								
11	163,33	5,10	9,66								
12	158,33	5,33	10,13								
13	173,33	5,30	10,52								
14	159,00	4,00	9,48								
15	138,33	3,77	10,14								
16	169,33	6,42	9,09								
17	147,33	5,00	9,50								
18	144,67	5,15	9,89								
19	133,33	3,23	9,85								
20	168,33	4,72	8,97								
21	129,33	2,92	10,23								

KUS – koraci u stranu (s)

Želimo vizualizirati podatke.

Rješenje: Frekvencijska tablica + vizualizacija:

STATISTICS → BASIC STATISTICS AND TABLES → FREQUENCY TABLES → ADVANCED



1. **ODABRATI VARIJABLE:**
VARIJABLE NA PRESKOK
OZNAČAVAMO – PRITISNUTI
<CTRL> TIPKU, TE S MIŠOM KLIKATI
NA ŽELJENE VARIJABLE
2. **ODABRATI ADVANCED:**
- ODABRATI 3 INTERVALA
3. **PRITISNITI FREQUENCY TABLES**

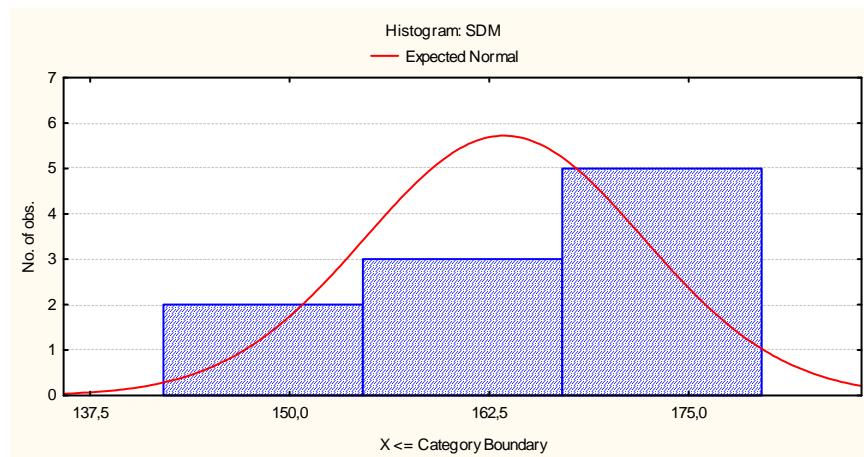
PREBACIVANJE DOBIVENE TABLICE/TABLICE PODATAKA U MS WORD

1. Smanjiti decimalna mjesta u dobivenoj tablici na dva decimalna mjesta.
2. Označiti cijelu tablicu (<CTRL> + C → COPY ili EDIT→COPY)
3. Otvoriti MS WORD program, START→ ALL PROGRAMS→ MICROSOFT OFFICE→ WORD
4. Pritisnuti <CTRL> + V → PASTE, tj. ZALJEPI ili EDIT→ PASTE

1. Insert → Object→ Create from File. Zatim odabrati BROWSE te pronaći file. (Za tablicu podataka)

Frequency table: KUS (vježba1.sta)				
	Count	Cumulative	Percent	Cumulative
8,476667<x<=9,170000	2	2	20,00000	20,0000
9,170000<x<=9,863333	7	9	70,00000	90,0000
9,863333<x<=10,55667	1	10	10,00000	100,0000
Missing	0	10	0,00000	100,0000

Odabirom opcije Histogram generiramo:



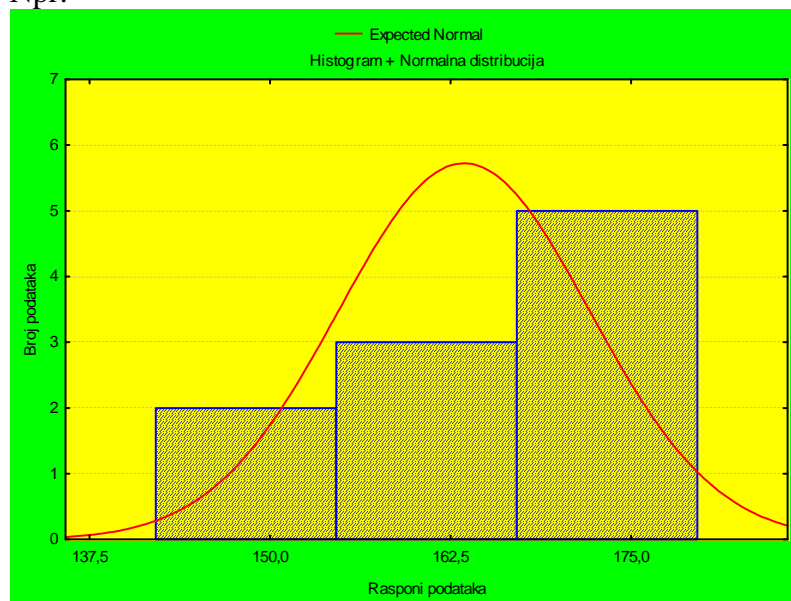
Parametri grafa se mogu editirati i u Wordu i u Statistici 7.0!

Word: Desni klik na grafu ->Graf Object -> Edit

Statistica: Desni klik na grafu -> Graph Properties

Obavezno promijeniti oznake na osima!!

Npr:



Parametri deskriptivne statistike:

Rezultati 10 ispitanika u tri motoričke varijable			
	1 SDM	2 BML	3 KUS
1	148,33	6,30	9,45
2	170,00	4,83	10,21
3	170,67	5,72	9,85
4	170,00	5,13	9,30
5	161,67	5,67	9,51
6	173,33	6,22	9,45
7	167,67	6,07	9,63
8	150,00	5,70	
9	159,33	6,58	
10	163,33	4,47	

ILI

Statistics -> Basic statistics and tables -> Advanced

ODREĐIVANJE DESKRIPTIVNIH PARAMETARA

STATISTICS → BASIC STATISTICS AND TABLES → DESCRIPTIVE STATISTICS

1. KARTICA ADVANCED

2. ODREDITI VARIABLE

3. OZNAČITI S KVAČICOM ONE DESKRIPTIVNE PARAMETRE KOJE ŽELIMO IZRAČUNATI

4. PRITISNUTI SUMMARY: DESCRIPTIVE STATISTICS

Dobivenu tablicu cijelu označiti i pritisnuti EDIT → COPY WITH HEADERS

Otvoriti MS WORD i pritisnuti EDIT → PASTE

Vježba: Oblikuj tablicu kao na slici!

Varijable	n	AS	Me	Mo	f_M	Min	Max	Q_1	Q_3	σ^2	σ	α_3	α_4
MKUS AS	31 2	9,94	9,85	Multipl e	^o	7,74	13,7	9,20	10,7	1,0	1,01	0,32	- 0,07
MPN AS	31 2	12,8 9	12,3 2	12	6	5,90	31,8	10,5 2	14,4	11,6	3,41	1,77	5,95
ISKRET AS	31 2	98,2 2	99,3 3	100	10	50,3 3	133, 3	89,3 3	107, 2	173, 8	13,1 8	- 0,35	0,22
M20 AS	31 2	4,08	4,07	Multipl e		2,85	5,0	3,88	4,3	0,1	0,33	0,13	0,61

NAPOMENA: Imena varijabli (MKUS AS, MPN AS, ISKRET AS, M20 AS, broj entiteta te izračunatri podaci su proizvoljni)

n – broj ispitanika (**Valid N**)

AS – aritmetička sredina (prosječni rezultat svih ispitanika) (**Mean**)

Me – medijan, rezultat ispitanika koji dijeli niz svih rezultata poredanih po redu na dva jednaka dijela (**Median**)

Mo – mod, rezultat ispitanika koji je najčešće postizan (**Mode**)

f_{M_0} – frekvencija moda (**Frequency of mode**)

Min – minimalni rezultat svih ispitanika (**Minimum**)

Max – Maksimalni rezultat svih ispitanika (**Maximum**)

Q_1 – donji kvartil, rezultat koji dijeli niz svih rezultata na $\frac{1}{4}$ i $\frac{3}{4}$ (**Lower quartile**)

Q_3 – gornji kvartil, rezultat koji dijeli niz svih rezultata na $\frac{3}{4}$ i $\frac{1}{4}$ (**Upper quartile**)

σ^2 – varijanca, srednje kvadratno odstupanje rezultata od prosječnog (aritmetičke sredine) (**variance**)

σ – standardna devijacija, mjera varijabiliteta (**standard deviation**)

α_3 – koeficijent asimetrije krivulje (**skewness**)

α_4 – koeficijent zakrivljenosti krivulje (**kurtosis**)

Napomena: sigma (σ) i alfa (α) se dobiju klikom na INSERT → SYMBOL → Basic Greek

Tko je najbolji???

Normalizacijom podataka te generiranjem nove varijable koja je aritmetička sredina varijabli dobivamo rješenje!!

Rezultati 10 ispitanika u tri motoričke varijable			
	1 SDM	2 BML	3 KUS
1	148,33	6,30	9,45
2	170,00	4,83	10,21
3	170,67	5,72	9,85
4	170,00	5,13	9,30
5	161,67	5,67	9,51
6	173,33		
7	167,67		
8	150,00		
9	159,33		
10	163,33		

Varijable kod kojih je veći rezultat entiteta znači lošiji rang, množimo sa (-1)
(Npr trčanje na 100 metara, Plivanje leđnim stilom na 50 metara...)

	Rezultati 10 ispitanika u tri motoričke varijable		
	1 SDM	2 BML	3 KUS
1	-1,73	0,93	-0,12
2	0,75	-1,23	1,81
3	0,83	0,07	0,89
4	0,75	-0,79	-0,49
5	-0,20	-0,00	0,04
6	1,14	0,81	-0,12
7	0,49	0,59	0,33
8	-1,54	0,05	0,53
9	-0,47	1,35	-1,71
10	-0,01	-1,77	-1,15

Desni klik na ime varijable i dobijemo izbornik:

	1 SDM	2 BML	3 KUS
1	-1,73	0,93	-0,12
2	0,75	-1,23	1,81
3	0,83	0,07	0,89
4	0,75	-0,79	-0,49
5	-0,20	-0,00	0,04
6	1,14	0,81	-0,12
7	0,49	0,59	0,33
8	-1,54	0,05	0,53
9	-0,47	1,35	-1,71
10	-0,01	-1,77	-1,15

Add Variables

How many: 1 Use 0 in "After" field to insert before first variable.
 After: KUS Double-click on it or press F2 to select variable from list.

Name: AKUMUL Type: Double

MD code: -9999 Length: 8

Display format

- General
- Number
- Date
- Time
- Scientific
- Currency
- Percentage
- Fraction
- Custom

If values of the new variable are to be computed, and the data set is large, it saves time to add variables and simultaneously recalculate their values using the Batch Transformations option (Data menu).

Long name (label or formula with Functions):

$=(v1+v2+v3)/3$

Formulas: use variable names or v1, v2, ..., v0 is case #.
 Examples: (a) = mean(v1:v3, sqrt(v7), AGE) (b) = v1+v2: comment (after:)

VAŽNO:

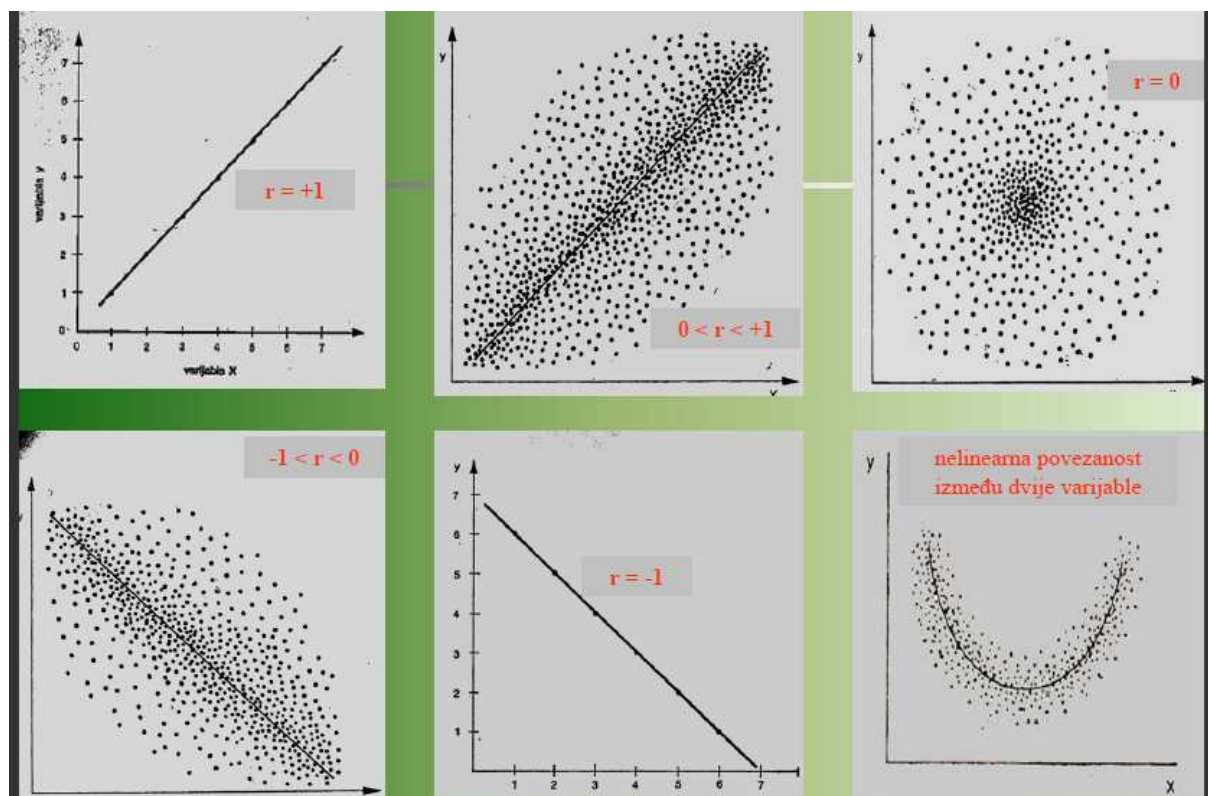
Varijabla koja akumulira rezultate, AKUMUL definirana je formulom $=(v1+v2+v3)/3$

Rezultati 10 ispitanika u tri motoričke varijable				
	1 SDM	2 BML	3 KUS	4 AKUMUL
1	-1,73	0,93	-0,12	-0,30692
2	0,75	-1,23	1,81	0,444157
3	0,83	0,07	0,89	0,596087
4	0,75	-0,79	-0,49	-0,17642
5	-0,20	-0,00	0,04	-0,05481
6	1,14	0,81	-0,12	0,608726
7	0,49	0,59	0,33	0,467736
8	-1,54	0,05	0,53	-0,32103
9	-0,47	1,35	-1,71	-0,27751
10	-0,01	-1,77	-1,15	-0,98003

Entitet sa najvećim rezultatom ima najbolji ukupni rezultat!

Linearna Regresija

Podsjetimo se mogućih odnosa dvaju varijabli:



VAŽNO!!!

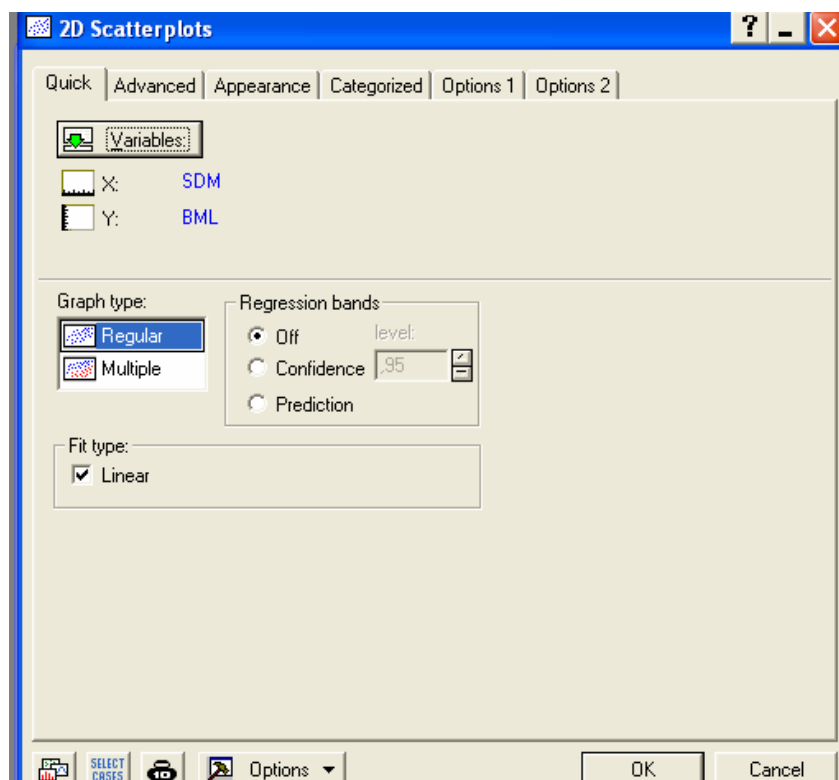
$r = 0.00$ do ± 0.20 → nikakva ili neznatna povezanost
 $r = 0.20$ do ± 0.40 → laka povezanost
 $r = 0.40$ do ± 0.70 → stvarna značajna povezanost
 $r = 0.70$ do ± 1.00 → visoka ili vrlo visoka povezanost

Visina korelacije nije samo odraz stupnja povezanosti, nego može biti i posljedica utjecaja drugih varijabli!!

Za detalje i Excel implementaciju pogledati:

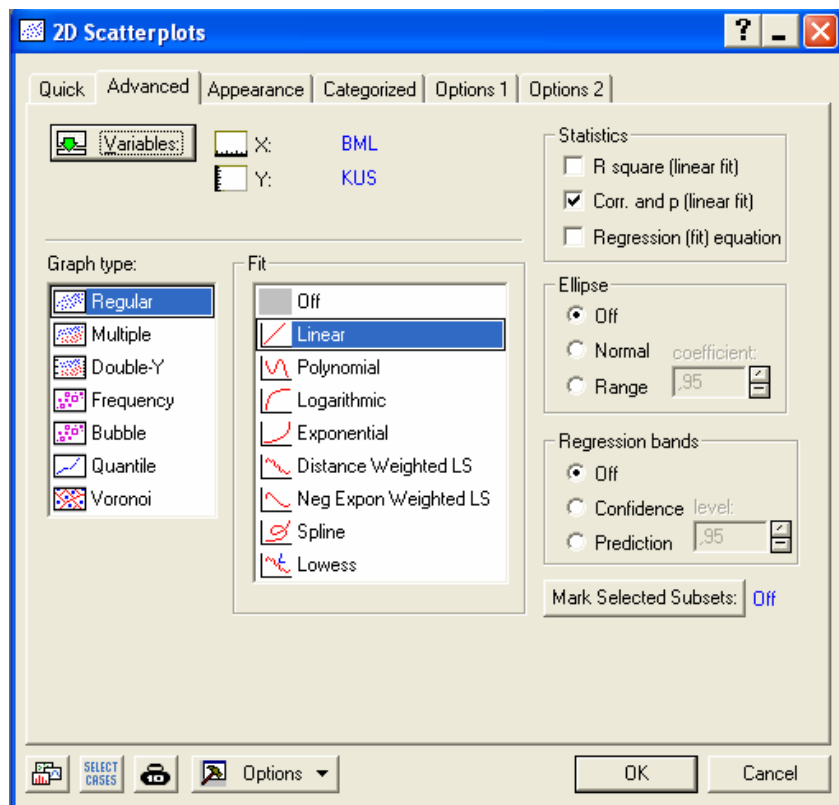
<http://dominis.phy.hr/~kvurnek/SAMP/SEMINAR/zavrsni.html>

Graphs-> Scatterplots ->

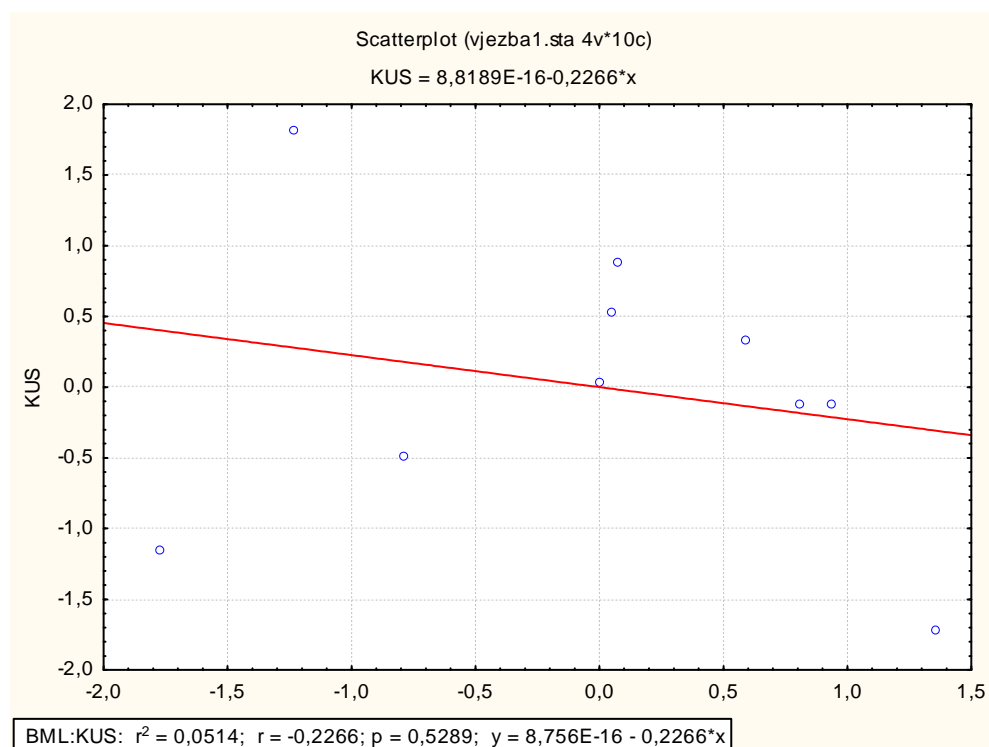


Klikom na Variables odabrati zavisnu (Y) i nezavisnu varijablu (X).

Odabirom Advanced biramo neke dodatne opcije:



Što uz stavljanje kvačice na R square, Corr. And p, Regression equation daje:

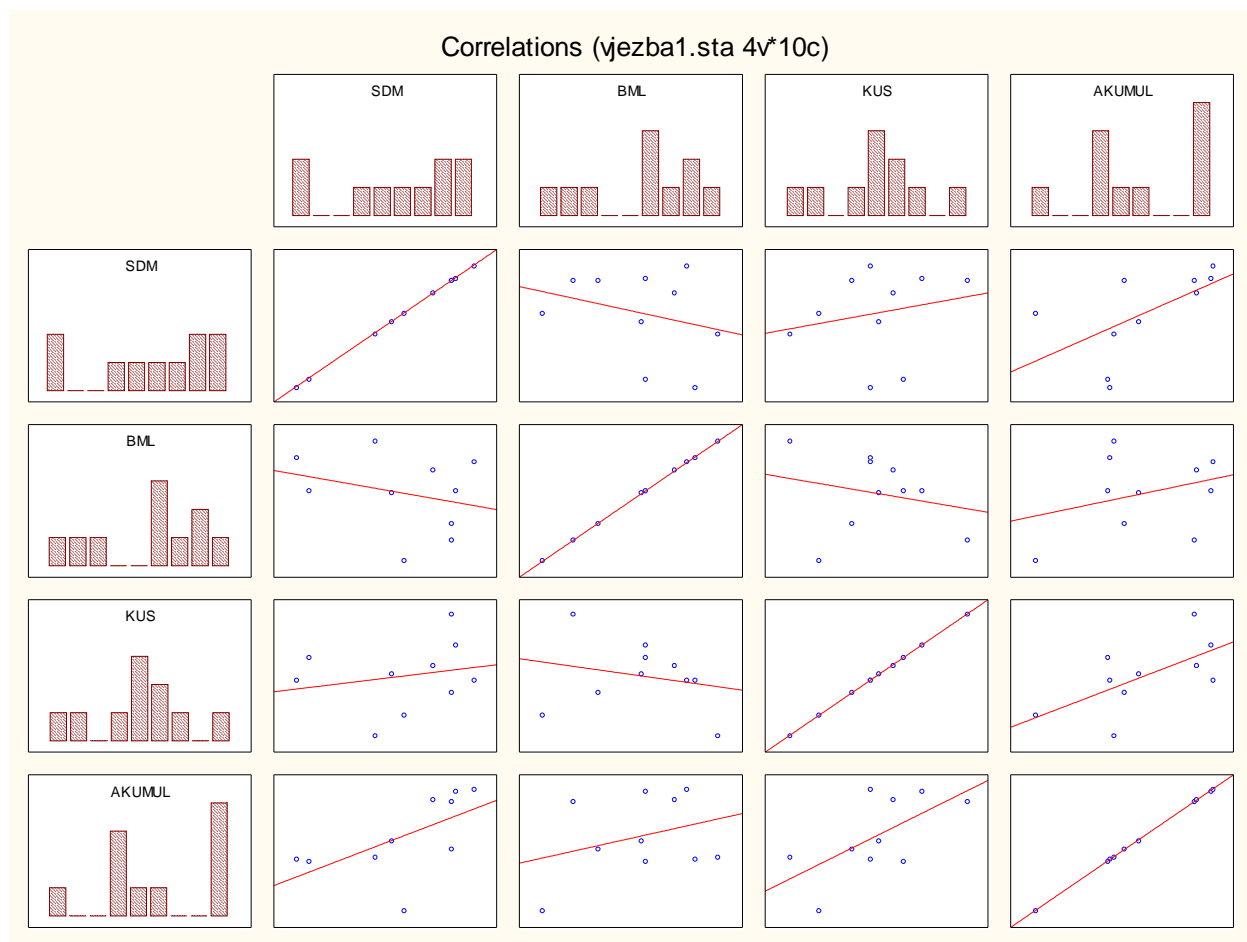


Uočimo da je linearna regresija specijalni slučaj multiple tj višestruke regresije za koju također postoji Statistica podrška.

Statistics->Multiple Regression

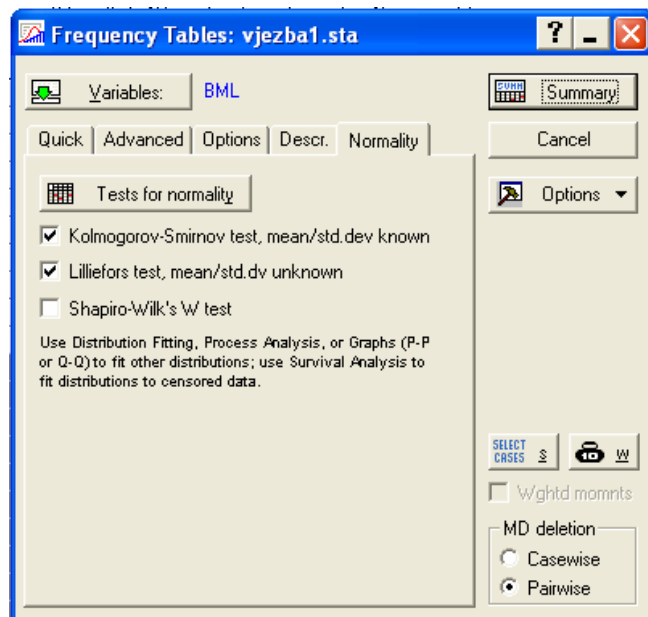
Odabirom:

Statistics->Basic Statistics/Tables -> Correlation Matrices ->Scatterplot Matrix -> Select All
(2 puta)
Generira se vizualizacija matrice korelacija.



K-S Test za testiranje normaliteta distribucije:

Statistics-> Basic Statistics/Tables ->Frequency Tables->



Što nakon odabira varijable BML daje:

Variable	N	max D	K-S p
BML	10	0,199018	p > .20

Max D – Maximalna razlika između opažene i očekivane frekvencije

K-S p – Vjerojatnost normaliteta distribucije

NAPOMENE:

Postoji više različitih načina realiziranja nekih proračuna Statisticom 7.0. Na usmenom ispitu ispitivati će se razumijevanje dotičnih proračuna.