Metode merenja i obrade podataka
Deskriptivna statistika

Šesto predavanje

Sadržaj

1. Šta je merenje
2. Varijable i konstante
3. Dizajn istraživanja i statistička analiza
4. Statističko zaključivanje
5. Organizacija podataka
6. Prikaz podataka
7. Mere centralne tendencije
8. Mere disperzije
9. Deskriptivna statistika u Excel-u

Šta je merenje

Šta je merenje (osnovni pojmovi)

MERENJE: Upoređivanje određene vrednosti sa zadatim (definisanim) standardom

PODATAK: Rezultat merenja

STATISTIKA: Skup matematičkih “tehnika” kojima se podaci organizuju, “tretiraju” i prikazuju za dalju interpretaciju i evaluaciju

EVALUACIJA: “Filozofski” koncept određivanja vrednosti, odnosno značaja dobijenih podataka

Osobine merenja

Svako merenje mora da bude precizno...

- **Validnost:**
  - Da li rezutat merenja u saglasnosti sa onim što bi trebalo da meri...
- **Pouzdanost:**
  - Mera ponovljivosti
- **Objektivnost:**
  - Uticaj različitih faktora izbegnut ili kontrolisan

Osobine merenja

Više o validnosti i pouzdanosi možete saznati na:

1. “ A New View of Statistics”

http://www.sportsci.org/recource/stats/index.html

Sportscience...
Home/Latest Issue
Contents by Issue
Research Resources
Sports Medicine
Sport Nutrition
Statistics
Tests/Technology
Train/Perform

Search...
This Site
The Web
Medline/Journals
Images: Google/Lycos
Other Searches

Links to...
Sportsci Email List
Other Email Lists
Jobs and Careers
Journal Homepages
Training Info
Universities
Other Links

About This Site...
Feedback
Info for Authors
Short Tour
Help!

SPORTSCIENCE

A Peer-Reviewed Site for Sport Research

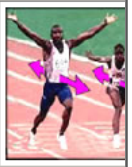
Volume 8

Peak Performance Online: top of the training sites

Journal of Sports Science and Medicine: on-line journal alternative

Visit the Sport Science Library at GSSI

Monitoring performance with tests: article



News and Comment

In Brief — Editorial: Page Numbers for Sportscience. Easier to cite
Clinical Significance and Decisiveness. Make better infe
Updated Endnote Journal Abbreviations. Less errors in r
Calibrating Metabolic Carts. Save on calgas. Nov 28; upd

Research Resources — Impact Factors of Journals in Sport and Exercise Scien
Will Hopkins. An assessment of the latest ratings. Nov 25

Perspectives

Tests/ Technology — How to Interpret Changes in an Athletic Performance Te
Article/slideshow on smallest worthwhile changes, the b
making sense of the results. Nov 25
Commentary. *Christopher Gore*. Nov 25
Commentary. *David Pyne*. Nov 25

Research Resources — An Introduction to Meta-analysis. *Will Hopkins*. Article/s
quantitative systematic reviews of original research. Nov
Bias in Bland-Altman but not Regression Validity Analy
Bland-Altman plots are misleading for method comparis
Commentary. *Alan Batterham*. Nov 29

Review

Train/ Perform — Effects of High-intensity Training on Performance and P
Endurance Athletes. *Carl Paton and Will Hopkins*. Gains
forms of resistance training. Nov 26
Commentary. *Carl Foster*. Nov 26
Commentary. *Philo Saunders and David Pyne*. Dec 2

SIRC index Sportscience articles in SPORTDiscus.

- Download Endnote file for this edition.
- Recent issues: 2003 · 2002 · Archive
- Become a contributor

Mail to the sportsci.org domain has been disabled, owing to volume of spam.

To contact the editor, click here .

Homepage · ©2004

New View of Stats: Home
Sportscience: Home

About These Pages
Alpha Reliability
ANCOVA (Analysis of Covariance)
ANOVA (Analysis of Variance)
One-Way
Two-Way
Repeated Measures
Arcsine-root Transformation
Assessing an Individual (using reli

Bayesian analysis
Become a license holder
Between-subject variation
Bias
Binomial regression
Bonferroni adjustment
Bootstrapping

Calibration equation
Categorical modeling
Central limit theorem
Chi-squared test
Clinical significance
Cluster analysis
Coefficient of variation
Defined
Measure of reliability
from log-transformed data
Use of ± and ×/÷
Complex Models
Confidence limits/Interval
Calculations: assumptions
of Correlation
Defined
of Effect Size (long, pop SD)
of Effect Size (long, sample SD)
of Effect Size (x-sect)
Example
Effect of Sample Size
of Frequency Difference
of Goodness-of-Fit
from a P Value
Powerpoint Presentation
Spreadsheet
Statistical Significance
Contingency Table
Contrasts: see Estimates
Controlled trial
Confounders
Controlling for something
Controlling Type I Error
Correlation Coefficient
Counts as dependent variable
Covariance Defined
Covariates in repeated measures
Crossovers
Simple
Multiple
Cumulative Type I and II error

Data and variables

A New View of Statistics

Will G Hopkins © 2004

Go to: Preface · Contents · Start of Text · Search · Home

This image has active links:

UPDATE **June 25:** Validity spreadsheet now has confidence limits for parameters of calibration equation. **June 2:** Corrected residuals in log section of validity spreadsheet. For history of **previous updates**, see below.

New original approaches to statistics for researchers: the examples are taken from exercise and sport science, but the principles apply to all empirical sciences. Read more in the preface.

Feedback wanted: if you can't understand something here, it's my fault. Email me.

Become a license holder: for a copy of this text for your own computer or for a local server, read this information.

Full Contents

Short Contents:

Reference: Hopkins, W. G. (2000). A new view of statistics. Internet Society for Sport Science: http://www.sportsci.org/resource/stats/.

If you are viewing this page without a navigation/index frame on the left, click here.

Go to: Preface · Contents · Start of Text · Search · Home

UPDATES

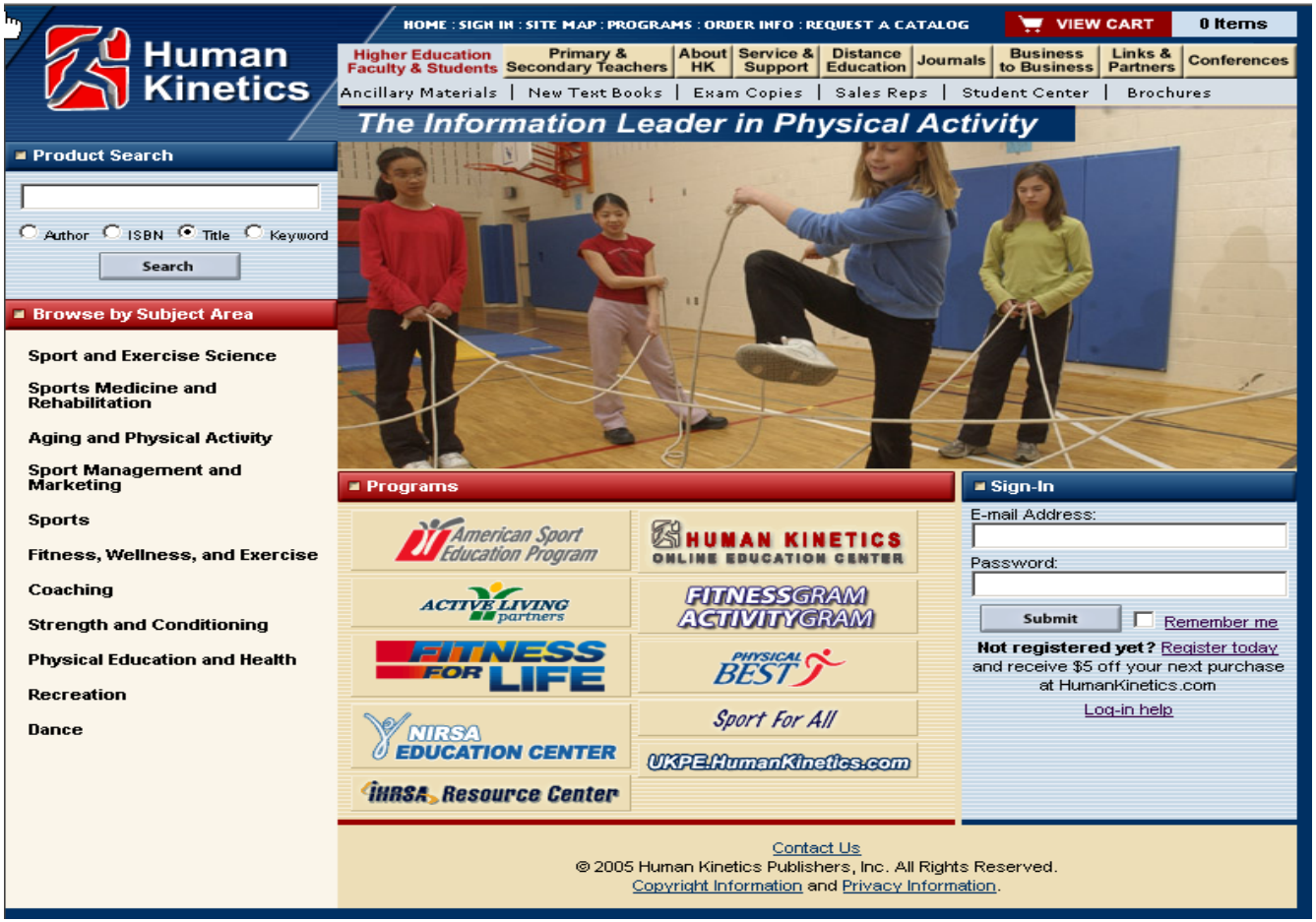
2004

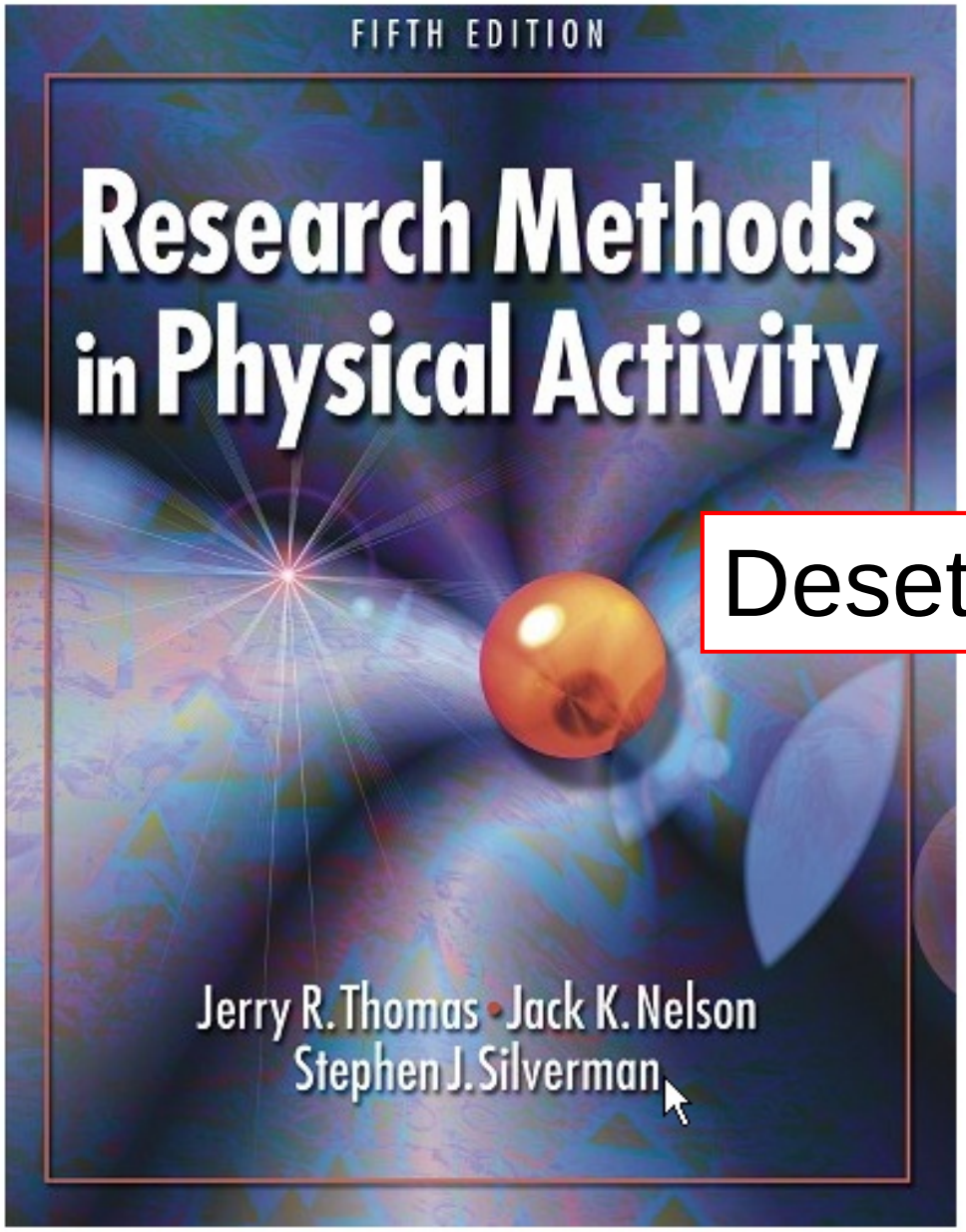
Aug 19: Info about generalized linear modeling for variables representing counts and proportions.

2003

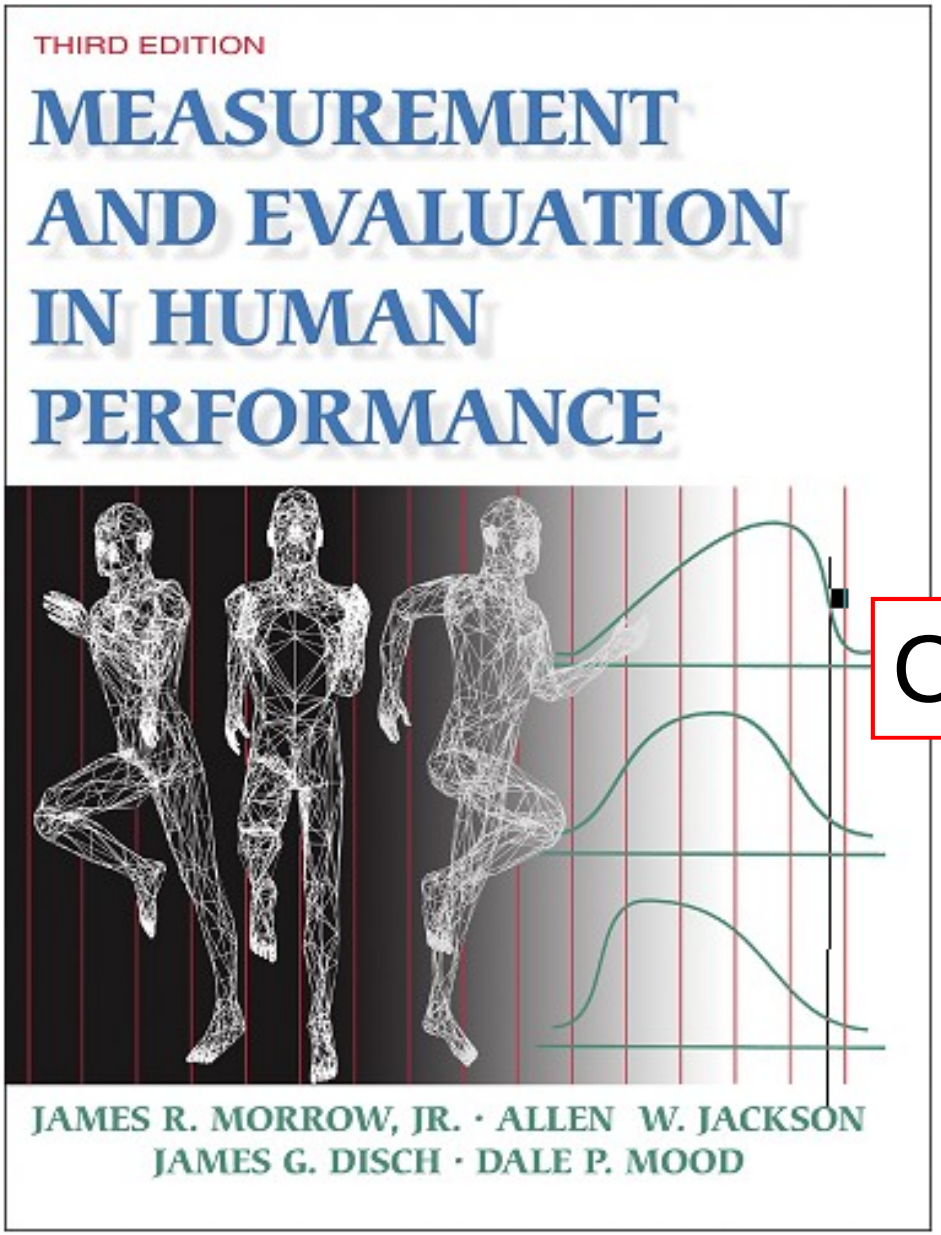
Nov 1: Article and spreadsheets for analysis of straightforward controlled trials, crossovers, and time series.
Oct 10: Adjustable level for confidence limits added to reliability and validity spreadsheets.
June 25: Minor improvement to Mean ± SD or Mean ± SEM?
June 15: Correcting change scores and other updates of regression to the mean.
June 8: Slideshow on repeated measures. Tweaking of categorical modeling and most pages on repeated measures. New section on repeated measures with troublesome variables.

http://www.humankinetics.com





Deseto poglavlje


THIRD EDITION
MEASUREMENT AND EVALUATION IN HUMAN PERFORMANCE
JAMES R. MORROW, JR. · ALLEN W. JACKSON
JAMES G. DISCH · DALE P. MOOD

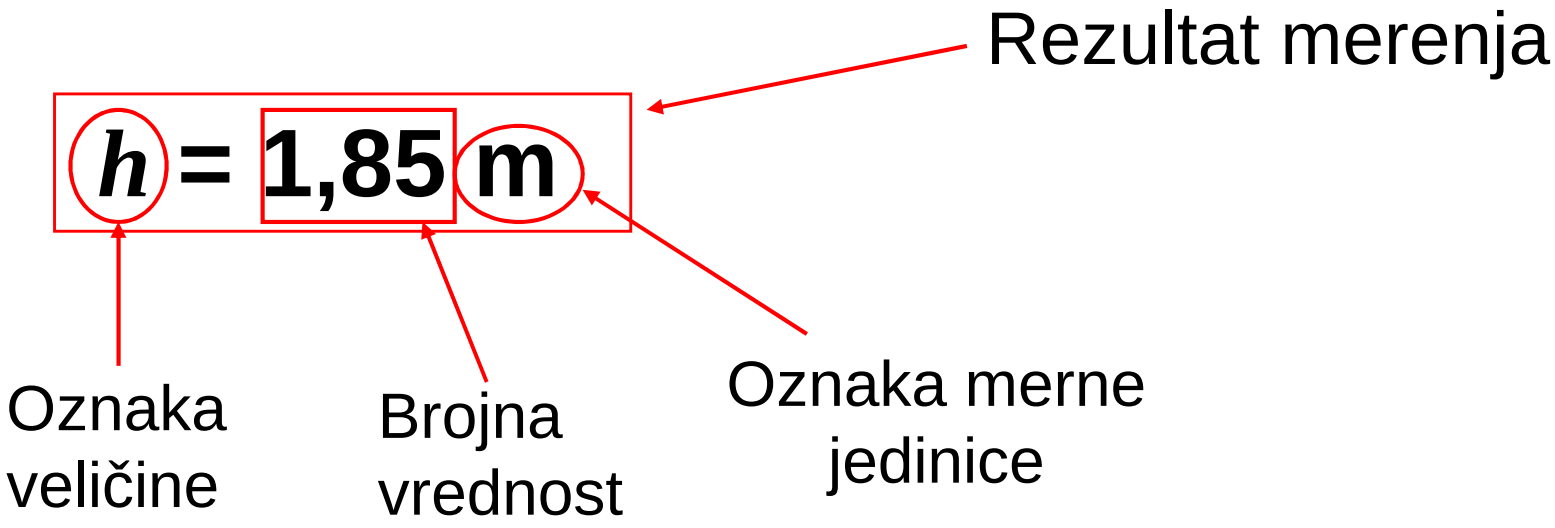
Cela knjiga

Merni postupak

Merni postupak

- Identifikacija objekta koji treba izmeriti
- Standard (jedinica mere)
- Proces upoređivanja (MERENJE!)...
- Kvantitativni zaključak...

Merni postupak



Kada rezultat merenja pridodamo odgovarajućoj ljudskoj osobini koju smo merili (recimo visini čoveka) rezultat "postaje" **varijabla (promenljiva)** **vidi nastavak...**

Varijable i konstante

Varijable i konstante

- **Varijabla** je karakteristika osobe, mesta, stvari ili procesa (dešavanja) koja može da ima više različtih vrednosti (promenljive)
- **Konstante (parametri)** su karakteristike koje se vremenom ne menjaju (nepromenljive)

Vrste i klasifikacija podataka

Varijable: Kontinualne i diskretne

Rezultati merenja odgovarajućih varijabli mogu se klasifikovati na više načina:

Prema objektivnosti merenja:

- **Kvantitativni rezultati (podaci)**
- **Kvalitativni rezultati (podaci)**

Prema skali merenja:

- **Nominalni (koje se prebrojavaju)**
- **Ordinalni (redosled)**
- **Intervalni (mogu imati negativne vrednosti)**
- **Racionalni (ne mogu biti negativne)**

Istraživački dizajn i statistička analiza

Testiranje hipoteze:

- Istraživačka hipoteza (H_n)
- Nulta hipoteza (H_0)

Ukoliko je H_0 tačna, H_n je netačna i obrnuto...

Nezavisne i zavisne promenljive

U zavisnosti od “mogućnosti” da na njih utičemo eksperimentalnim dizajnom...

- **Nezavisne (prediktorske)**
- **Zavisne (kriterijumske)**

Validnost eksperimenta

Eksperiment (kao deo istraživačkog dizajna) mora da poseduje i tzv. **“unutrašnju”** (internal) i tzv. **“spoljašnju”** (external) **validnost**.

Zaključivanje u statistici

Zaključivanje u statistici

- **Populacija**: ma koja grupa pojedinaca, mesta ili stvari koje imaju bar jednu zajedničku osobinu
- **Uzorak:** deo populacije, koji je predmet statističke “obrade”

Greška predviđanja je obrnuto srazmerna veličini uzorka

Odabir uzorka

- **Slučajnim odabirom**: svaki član populacije ima jednake šanse da bude izabran

- **Stratifikovano "uzorkovanje":** prethodno populaciju delimo u odgovarajuće grupe (koje imaju nešto zajedničko...)

Odabir uzorka

Ukupan broj studentata	1000
Uzorak	50
Uzorak (%)	5.00%

	I godina	II godina	III godina	IV godina
Broj studenata po godinama	400	250	200	150
Uzorak	20	13	10	8

Parametri i statistika

Parametar

– karakteristika čitave populacije

Statistika

– Karakteristika uzorka

Parametri i statistika

Svaka procena parametra na osnovu statistike uzorka ima izvesnu “grešku”

Vrednost “greške” se nikada na zna pouzdano ali se može proceniti na osnovu veličine i varijabiliteta uzorka

Prikaz podataka

Raspodele

- Prikaz po redosledu
- Raspodela po frekvencijama
- Raspodela po grupnim frekvencijama

U zavisnosti od vrste podataka:

- Tabelarno
- Grafički

Organizovanje podataka

Opseg (*R*): Najveća vrednost (*H*) manje najmanja vrednost (*L*):

$$R = H - L$$

$$R = H - L + 1^{*}$$

* *Ukoliko se uračunaju i vrednosti na "krajevima"*

Prikaz po redosledu

Primer: Prikazani su rezultati testiranja 15 dečaka (zgibovi sa dlanovima okrenutim ka “spolja”):

12, 10, 9, 8, 2, 5, 18, 15, 14, 17, 13, 12, 8, 9, 16

Prikaz po redosledu

Table 2.1 Rank Order Distribution of Pull-Up Scores

X
18
17
16
15
14
13
12
12
10
9
9
8
8
5
2

$N = 15$
$H = 18$
$L = 2$
$R = 18 - 2 = 16$

Raspodela po frekvencijama

Table 2.2 Simple Frequency Distribution of Pull-Up Scores

X	f
20	2
19	0
18	3
17	6
16	8
15	10
14	17
13	21
12	25
11	24
10	26
9	19
8	16
7	12
6	10
5	4
4	3
3	2
2	1
1	2
0	1
	212

$N = 212$
$H = 20$
$L = 0$
$R = 20 - 0 = 20$

Raspodela grupnih frekvencija

Table 2.3 Grouped Frequency Distribution: Mile-Run Times in Seconds

X	f
580-599	3
560-579	9
540-559	13
520-539	15
500-519	17
480-499	21
460-479	19
440-459	25
420-439	23
400-419	18
380-399	15
360-379	12
340-359	9
320-339	5
300-319	2
	$N = 206$

Interval = Opseg/15

Raspodela grupnih frekvencija

Table 2.4 Grouped Frequency Distribution: Times in Seconds for 100-Meter Swim

X	f
115-119	1
110-114	2
105-109	3
100-104	7
95-99	10
90-94	15
85-89	17
80-84	11
75-79	5
70-74	4
65-69	2
60-64	2
	$N = 79$

Histogram

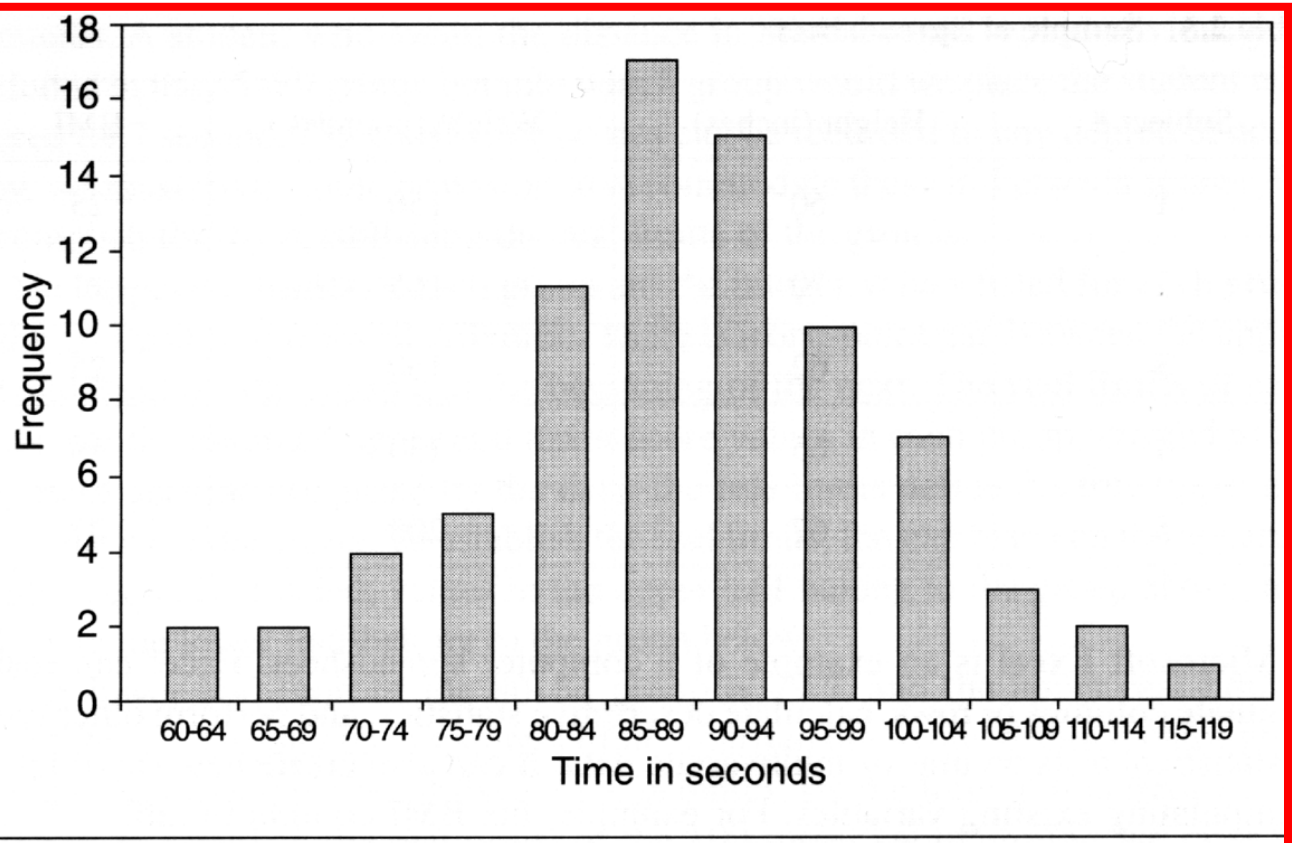
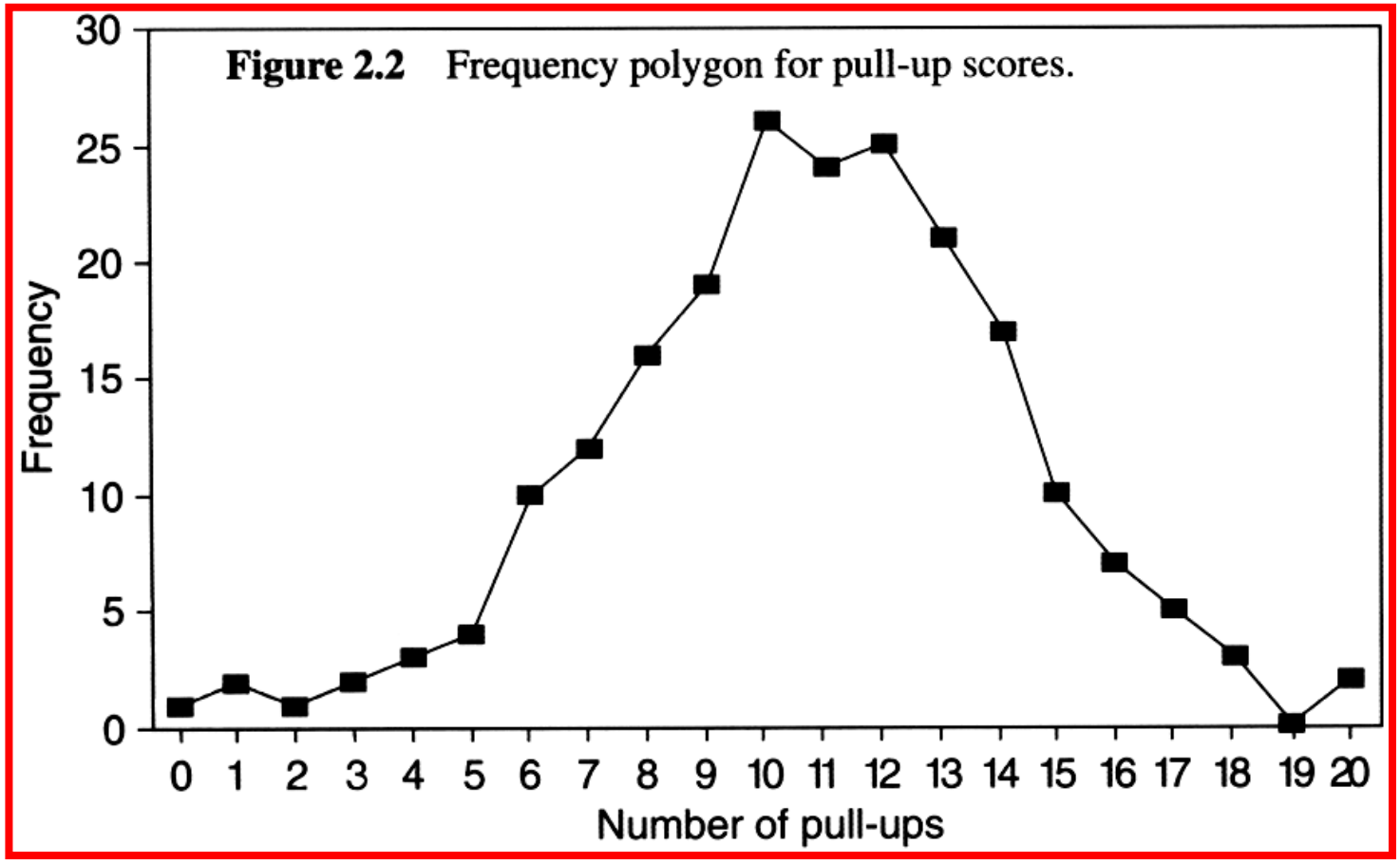


Figure 2.1 Histogram of 100-meter swim.

Poligon frekvencija



MERE CENTRALNE TENDENCIJE

Sadržaj

1. Mere centralne tendencije
2. Mere disperzije
3. Deskriptivna statistika u Excel-u

Mere centralne tendencije

MEDIJANA (Centralna vrednost)

MODUS (Najčešća vrednost)

SREDNJA VREDNOST (Aritmetička sredina)

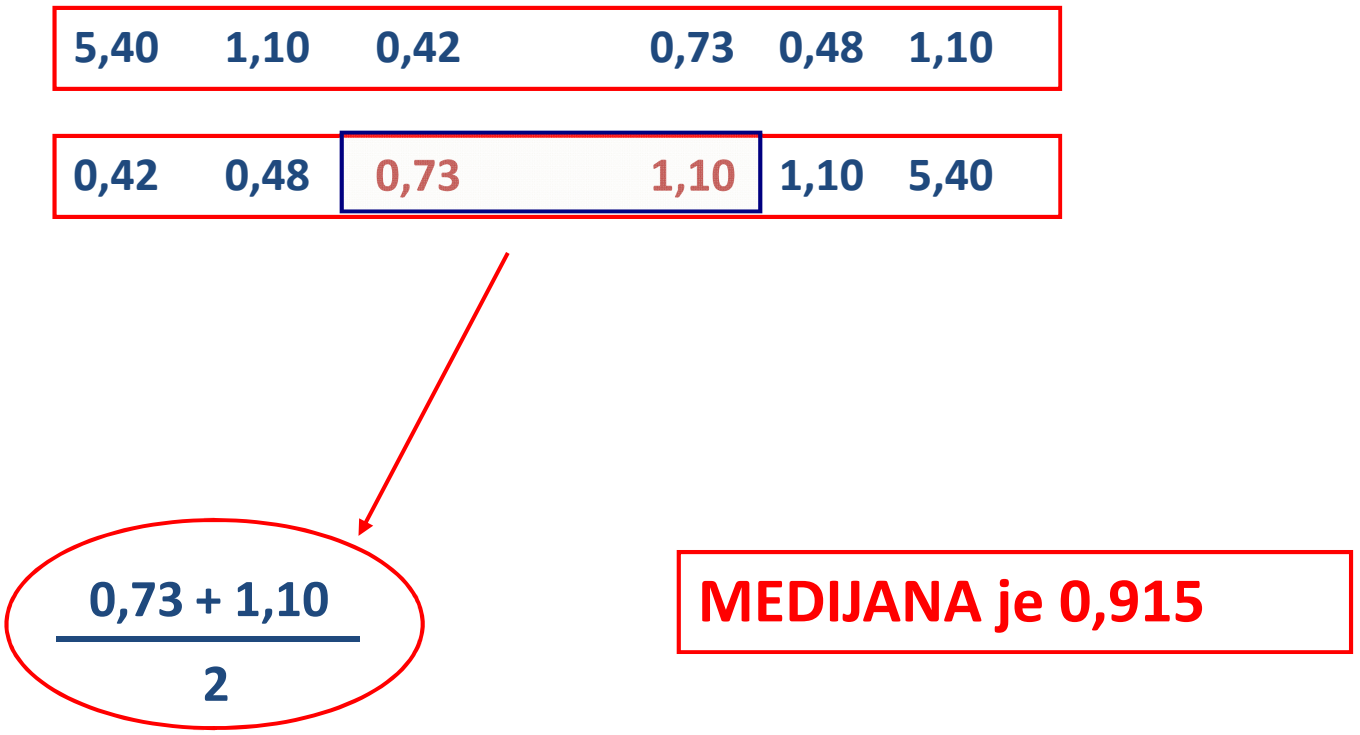
Medijana

Podatke poređaj po rastućem redosledu:

Odredi položaj (C) (koji je po redu) centralnog podatka: $C = (N+1)/2$

- za neparan broj podataka na tom *(“C-tom”)* položaju se nalazi medijana.
- za paran broj podataka dva su rezutata u sredini pa je medijana srednja vrednost ta dva “centralna” podatka

Medijana (paran broj podataka)



Medijana (neparan broj podataka)

5,40	1,10	0,42		0,73	0,48	1,10	0,66

0,42	0,48	0,66	0,73	1,10	1,10	5,40

MEDIJANA je 0,73

Modus

a. 5.40 1.10 0.42 0.73 0.48 1.10 ⇦ Modus je 1.10

b. 27 27 27 55 55 55 88 88 99 ⇦ Dvostruki modus - 27 & 55

c. 1 2 3 6 7 8 9 10 ⇦ Nema modusa

Aritmetička sredina

Zbir svih podataka podeli brojem podataka

$$\bar{x} = \frac{\sum x_i}{N}$$

x_i - *"i-ti" podatak,* N*-ukupan broj podataka*

Zajednička aritmetička sredina

Zbir proizvoda srednjih vrednosti podataka i njihovog broja podeli ukupnim brojem svih podataka

$$\bar{x} = \frac{\sum N_i \bar{x}_i}{\sum N_i}$$

$N_i \bar{x}_i$ - *proizvod "i-te" srednje vrednosti*

i broja podataka iz kojeg je ta srednja vrednost izračunata

Aritmetička sredina (raspodela podataka prema učestanosti)

Sumu proizvoda učestanosti pojavljivanja i odgovarajućih vrednosti podeli ukupnim brojem podataka

$$\bar{x} = \frac{\sum f_i x_i}{N}$$

x_i - *"i-ti" podatak,*

f_i *-ukupan broj podataka*

$N = \sum f_i$

MERE DISPERZIJE

Opseg (raspon)

Opseg (*R*): Najveća vrednost (*H*) manje najmanja vrednost (*L*):

$$R = H - L$$

$$R = H - L + 1^{*}$$

* *Ukoliko se uračunaju i vrednosti na "krajevima"*

Kvartili

KVARTILI

1. Podaci se poređaju od najmanjeg do najvećeg.
2. Q_1 - Određujemo kao medijanu prvih 50% podataka.
3. Q_3 - Određujemo kao medijanu drugih 50%podataka.

Kvartili

Međukvartilni opseg:

$$I = Q_3 - Q_1$$

Srednje (absolutno) odstupanje

$$Srednje\ odstupanje = \frac{\sum \left| x_i - \bar{x} \right|}{N}$$

x_i - "i-ti" podatak
$\bar{x}$ – aritmetička sredina
N – broj podataka

Varijansa

$$\sigma^2 = \frac{\sum \left(x_i - \bar{x} \right)^2}{N - 1}$$

$$\sigma^2 = \frac{\sum x_i^{\,2} - \frac{\left(\sum x_i \right)^2}{N}}{N - 1}$$

- σ^2 – varijansa
- σ – standardna devijacija
- x_i - “i-ti” podatak
- $\bar{x}$ – aritmetička sredina
- N – broj podataka

Varijansa (raspodela podataka prema učestanosti)

$$\sigma^2 = \frac{\sum fx^2 - \frac{\left(\sum fx\right)^2}{N}}{N-1}$$

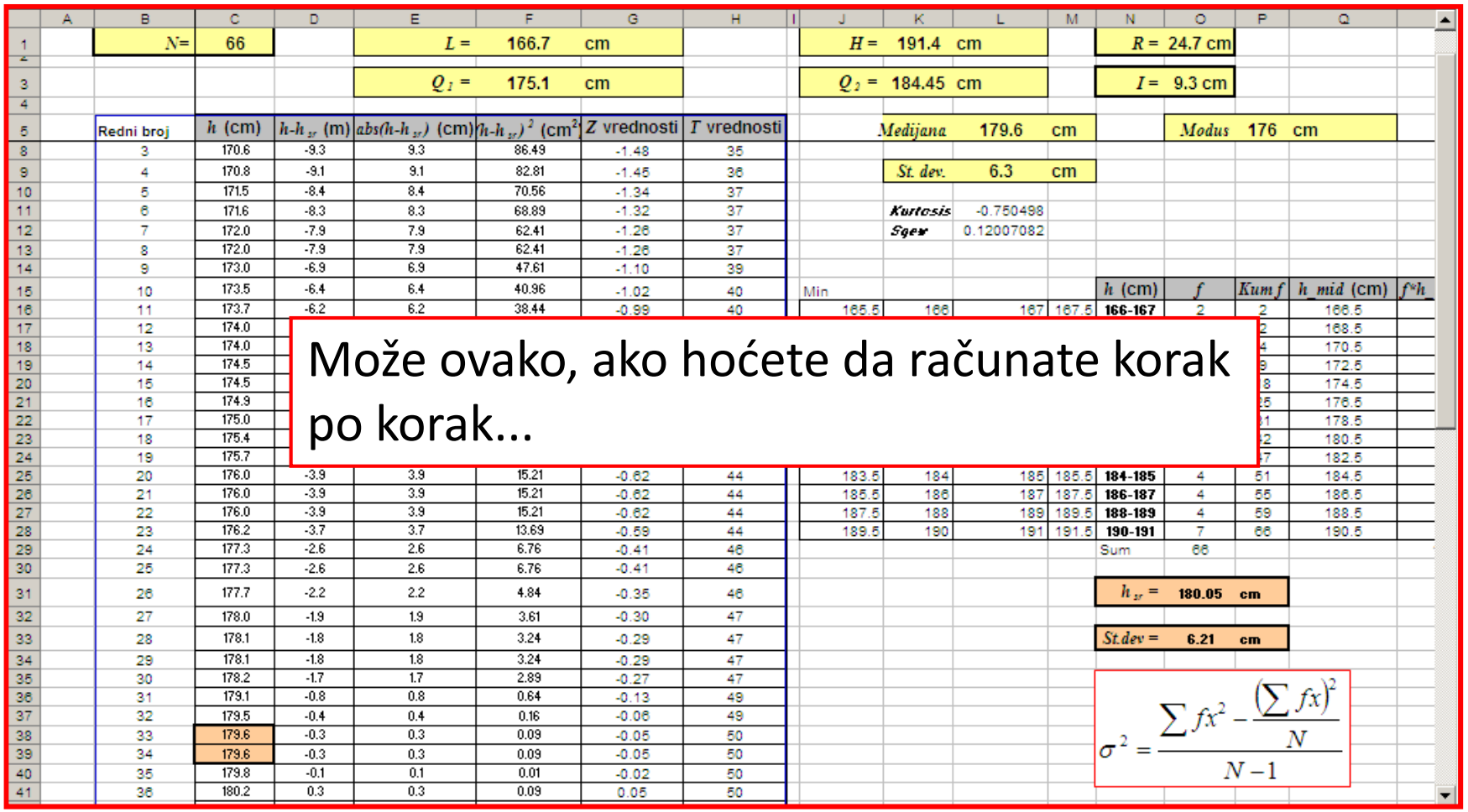
- σ^2 – varijansa
- σ – standardna devijacija
- x_i - “i-ti” podatak
- x – aritmetička sredina
- f_i – učestanost “i tog podatka

Standardna devijacija

Standardna devijacija: Kvadratni koren varijanse:

$$\sigma = \sqrt{\sigma^2}$$

Deskriptivna statistika u Excelu

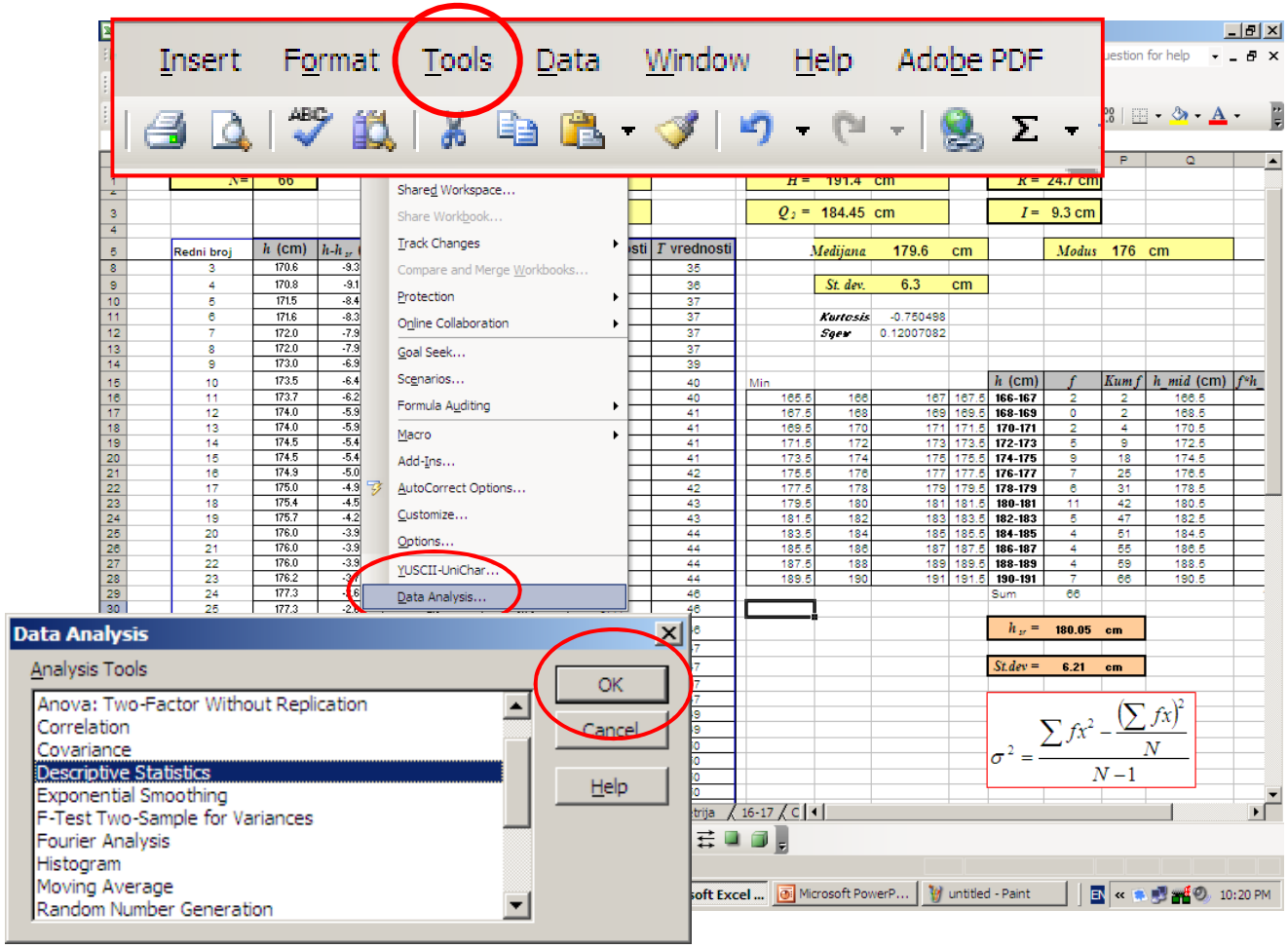


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1		N=	66		L =	166.7	cm			H =	191.4	cm		R =	24.7	cm	
3					Q_1 =	175.1	cm			Q_2 =	184.45	cm		I =	9.3	cm	
5		Redni broj	h (cm)	$h-h_{sr}$ (m)	$abs(h-h_{sr})$ (cm)	$(h-h_{sr})^2$ (cm^2)	Z vrednosti	T vrednosti		Medijana	179.6	cm		Modus	176	cm	
8		3	170.6	-9.3	9.3	86.49	-1.48	35									
9		4	170.8	-9.1	9.1	82.81	-1.45	36		St. dev.	6.3	cm					
10		5	171.5	-8.4	8.4	70.56	-1.34	37									
11		6	171.6	-8.3	8.3	68.89	-1.32	37		Kurtosis	-0.750498						
12		7	172.0	-7.9	7.9	62.41	-1.26	37		Sger	0.12007082						
13		8	172.0	-7.9	7.9	62.41	-1.26	37									
14		9	173.0	-6.9	6.9	47.61	-1.10	39									
15		10	173.5	-6.4	6.4	40.96	-1.02	40		Min				h (cm)	f	Kum f	h_mid (cm)	f*h_
16		11	173.7	-6.2	6.2	38.44	-0.99	40		165.5	166	167	167.5	166-167	2	2	166.5	
17		12	174.0													2	168.5	
18		13	174.0													4	170.5	
19		14	174.5													9	172.5	
20		15	174.5													8	174.5	
21		16	174.9													5	176.5	
22		17	175.0													1	178.5	
23		18	175.4													2	180.5	
24		19	175.7													7	182.5	
25		20	176.0	-3.9	3.9	15.21	-0.62	44		183.5	184	185	185.5	184-185	4	51	184.5	
26		21	176.0	-3.9	3.9	15.21	-0.62	44		185.5	186	187	187.5	186-187	4	55	186.5	
27		22	176.0	-3.9	3.9	15.21	-0.62	44		187.5	188	189	189.5	188-189	4	59	188.5	
28		23	176.2	-3.7	3.7	13.69	-0.59	44		189.5	190	191	191.5	190-191	7	66	190.5	
29		24	177.3	-2.6	2.6	6.76	-0.41	46						Sum	66		
30		25	177.3	-2.6	2.6	6.76	-0.41	46									
31		26	177.7	-2.2	2.2	4.84	-0.35	46						h_{sr} =	180.05	cm	
32		27	178.0	-1.9	1.9	3.61	-0.30	47									
33		28	178.1	-1.8	1.8	3.24	-0.29	47						St.dev =	6.21	cm	
34		29	178.1	-1.8	1.8	3.24	-0.29	47									
35		30	178.2	-1.7	1.7	2.89	-0.27	47									
36		31	179.1	-0.8	0.8	0.64	-0.13	49									
37		32	179.5	-0.4	0.4	0.16	-0.06	49									
38		33	179.6	-0.3	0.3	0.09	-0.05	50									
39		34	179.6	-0.3	0.3	0.09	-0.05	50									
40		35	179.8	-0.1	0.1	0.01	-0.02	50									
41		36	180.2	0.3	0.3	0.09	0.05	50									

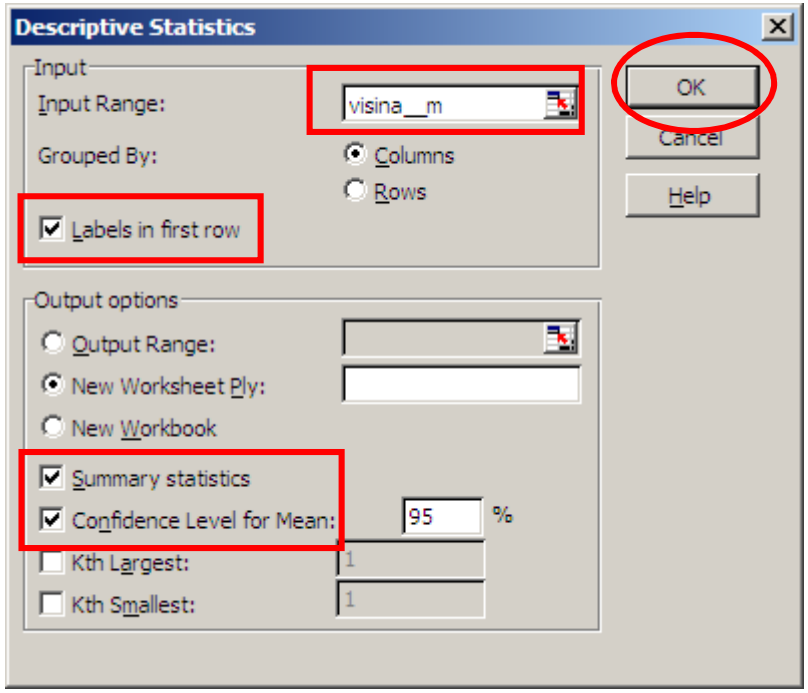
Može ovako, ako hoćete da računate korak po korak...

$$\sigma^2 = \frac{\sum fx^2 - \frac{\left(\sum fx\right)^2}{N}}{N-1}$$

Deskriptivna statistika u Excelu



Deskriptivna statistika u Excelu



	A	B
1	*h (cm)*	
2		
3	Mean	179.90
4	Standard Error	0.77
5	Median	179.6
6	Mode	176.0
7	Standard Deviation	6.3
8	Sample Variance	39.4
9	Kurtosis	-0.750
10	Skewness	0.120
11	Range	24.7
12	Minimum	166.7
13	Maximum	191.4
14	Sum	11873.4
15	Count	66
16	Confidence Level(95.0%)	1.5