



Statistička obrada podataka



v.as.mr. Samir Lemeš
slemes@mf.unze.ba



Tempus

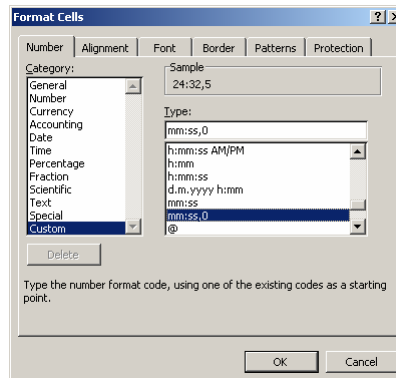


Statistička obrada podataka

- MS Excel kao alat za statističku obradu
- Standardne funkcije
- Analysis Toolpak
- Histogrami
- Crtanje grafikona
- Vremenski dijagrami
- Regresija i funkcionalna zavisnost
- ASCII tekst datoteke

Standardne funkcije

- =NOW()
- Format Cells
Number
Custom

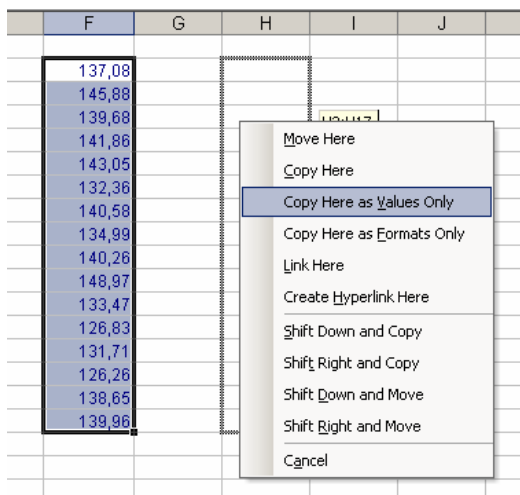


Standardne funkcije

- =RAND() ... generiše pseudoslučajne brojeve od 0 do 1
- Primjer 1: generisati slučajne brojeve između 125 i 150
- Rješenje: =125+25*RAND()
- Primjer 1: generisati slučajne brojeve između -1,5 i 0,5
- Rješenje: =-1,5+2*RAND()

Standardne funkcije

- Copy as Values Only



Standardne funkcije

- =AVERAGE(raspon)
Aritmetička sredina

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

- =STDEV(raspon)
Standardna devijacija uzorka

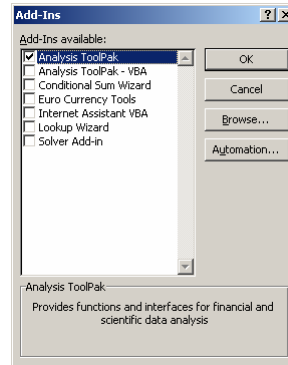
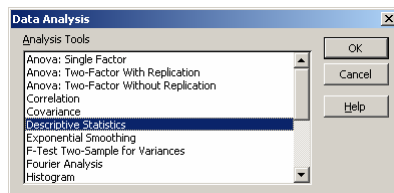
$$\sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

- =STDEVP(raspon)
Standardna devijacija populacije

$$\sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

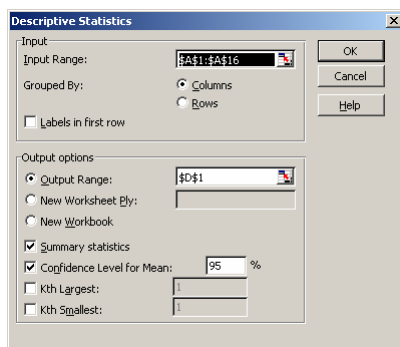
Analysis Toolpak

- Tools, Add-Ins...
- Tools, Data Analysis



Analysis Toolpak

- Tools, Data Analysis, Descriptive Statistics



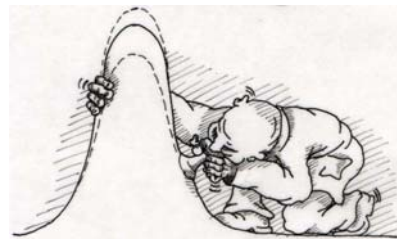
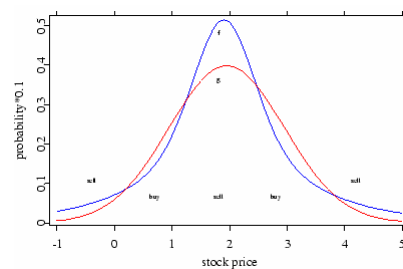
	A	B	C	D	E	F
1	13			Column1		
2	14					
3	13			Mean	13,5	
4	13			Standard Error	0,129099	
5	14			Median	13,5	
6	14			Mode	13	
7	14			Standard Deviation	0,516398	
8	14			Sample Variance	0,266667	
9	13			Kurtosis	-2,30769	
10	13			Skewness	0	
11	13			Range	1	
12	14			Minimum	13	
13	14			Maximum	14	
14	13			Sum	216	
15	13			Count	16	
16	14			Confidence Level(95,0%)	0,275169	
17						
18						

Analysis Toolpak

- **Mean:** aritmetička sredina
- **Standard Error:**
standardna devijacija distribucije / \sqrt{n}
(standardna greška procjene aritmetičke sredine)
- **Median:** vrijednost od koje lijevo i desno ima isti broj rezultata
- **Mode:** vrijednost koja se najčešće javlja ($f=\max$)
- **Standard Deviation:** $s = \text{STDEV}()$
- **Sample Variance:** s^2

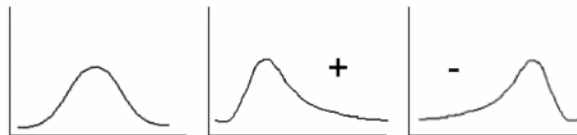
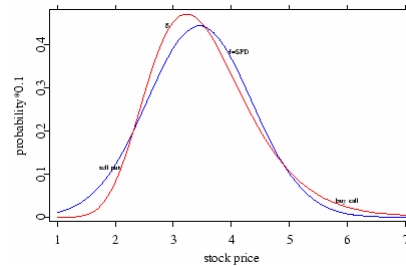
Analysis Toolpak

- **Kurtosis:**
zaobljenost vrha
krivulje distribucije
frekvencija
- Idealno: 0
- <0 : ravnija od normalne
- >0 : strmija od normalne



Analysis Toolpak

- **Skewness:**
asimetrija krivulje
distribucije frekvencija
- Idealno: 0
- <0 : vrh nagnut udesno
- >0 : vrh nagnut ulijevo



Analysis Toolpak

- **Range:** MAX-MIN
- **Minimum:** najveća vrijednost
- **Maximum:** najmanja vrijednost
- **Sum:** zbir svih vrijednosti
- **Count:** n
- **Confidence Level:**
vjerovatnoća rasipanja rezultata

Analysis Toolpak

- Prikaz rezultata: $x = \bar{x} \pm \lambda \cdot \frac{s}{\sqrt{n-1}}$

- λ : faktor pokrivanja

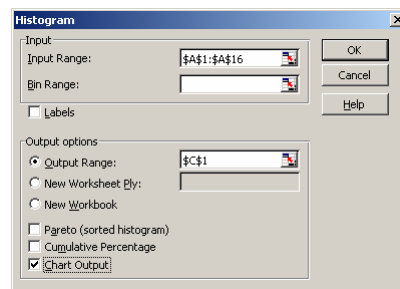
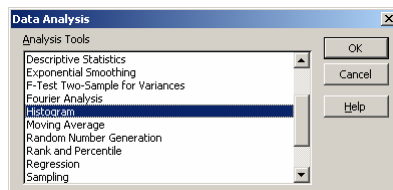
- $x = \text{Mean} \pm \text{Confidence Level}$ sa vjerovatnoćom od ___%

- $x = 13,50 \pm 0,28$ (P=95%)

Mean	13,5
Standard Error	0,129099
Median	13,5
Mode	13
Standard Deviation	0,516398
Sample Variance	0,266667
Kurtosis	-2,30769
Skewness	0
Range	1
Minimum	13
Maximum	14
Sum	216
Count	16
Confidence Level(95,0%)	0,275169

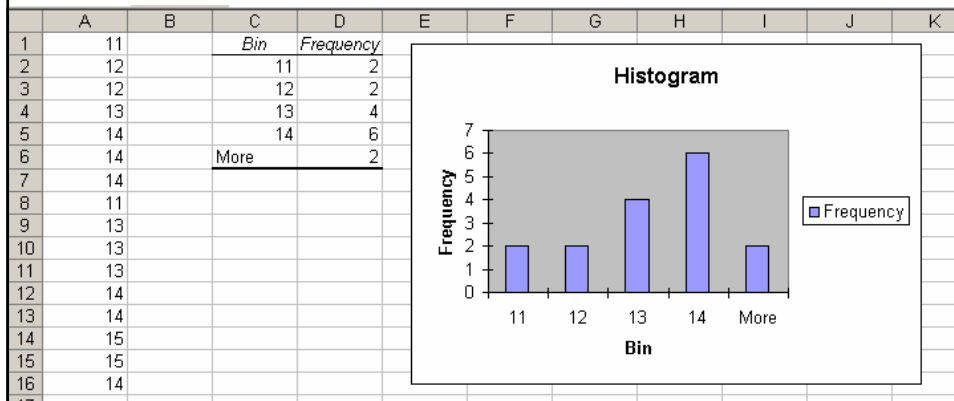
Histogramami

- Tools, Data Analysis, Histogram
- Input Range: podaci za analizu
- Bin: Range: vrijednosti za kategorije
- Chart Output



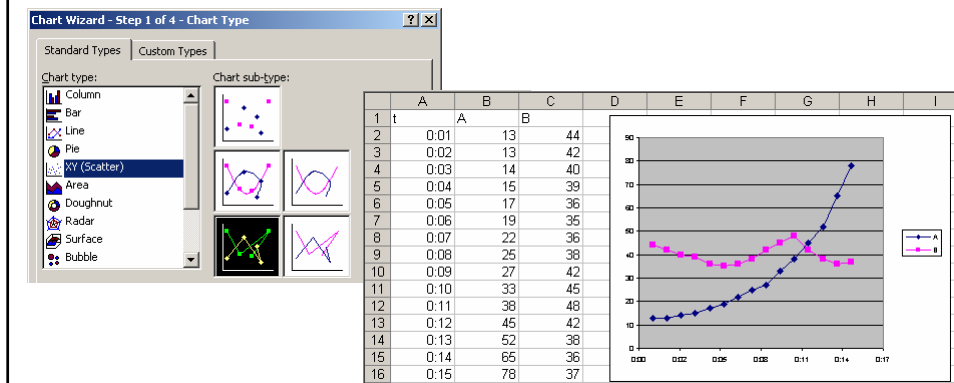
Histogrami

- Ako se ne navede "Bin Range", određuje se automatski



Crtanje grafikona

- XY Scatter – daje funkcionalnu zavisnost 2 ili više veličina

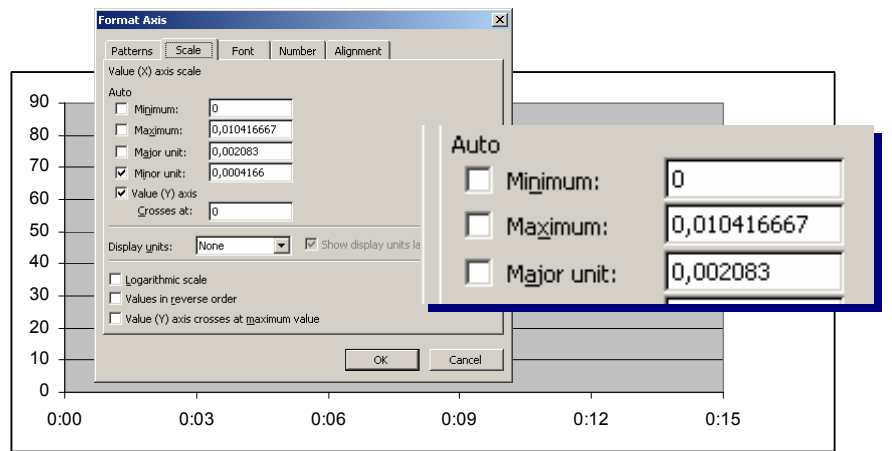


Vremenski dijagrami

- XY Scatter
- U prvoj koloni su vrijednosti prikazane kao vrijeme (sekunde, minute, sati, datumi)
- Vrijeme Excel prikazuje kao decimalne brojeve: 1,0 = 24h
- 1 sat = $1/24 = 0,04166667$
- 15' = $15 \cdot 1/(24 \cdot 60) = 0,010416667$
- 3' = $3 \cdot 1/(24 \cdot 60) = 0,002083333$

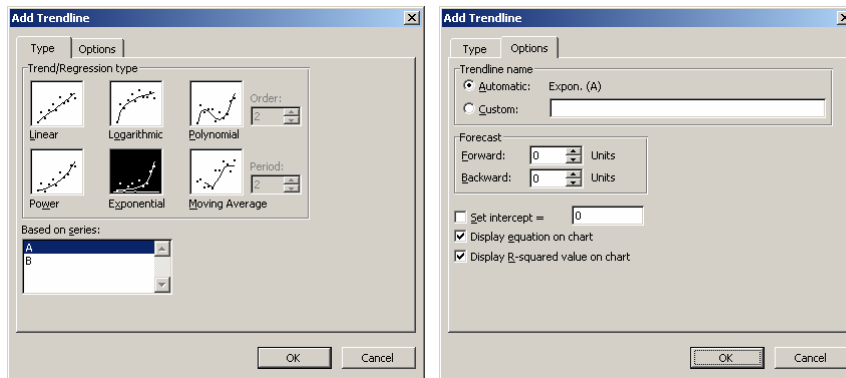
Vremenski dijagrami

- 15' = 0,010416667 3' = 0,002083333

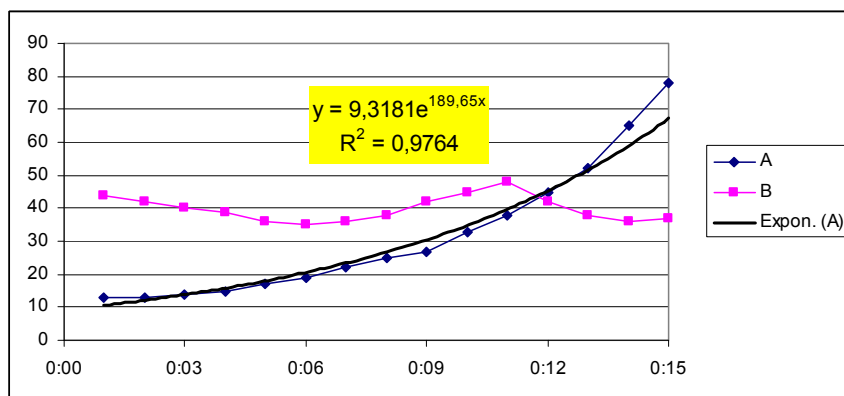


Regresija i funkcionalna zavisnost

- XY Scatter
- Add Trendline



Regresija i funkcionalna zavisnost



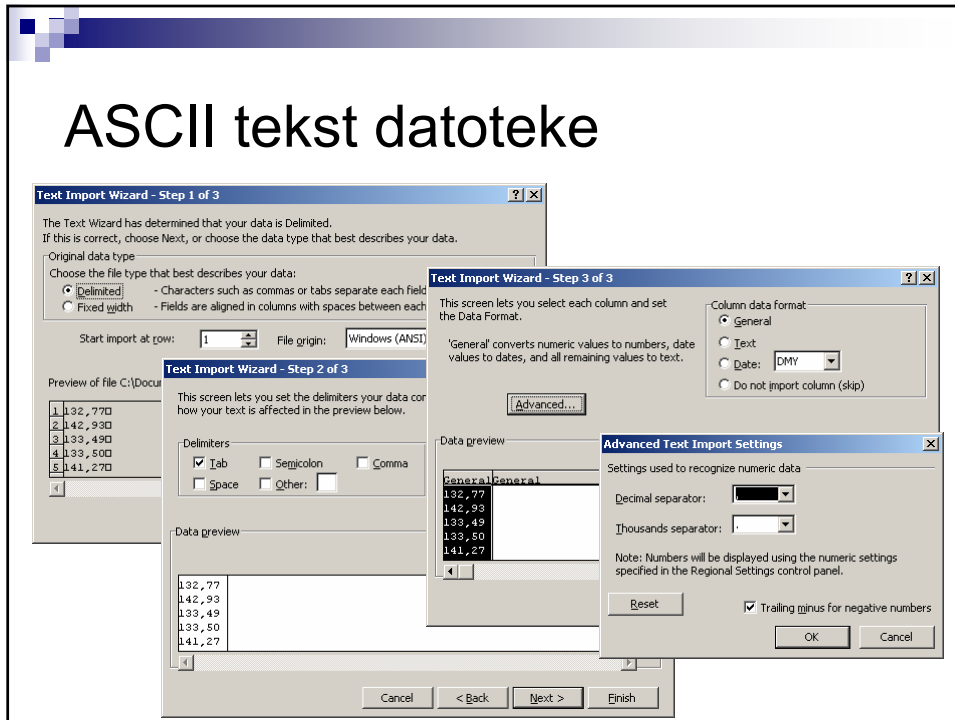
Regresija i funkcionalna zavisnost

- R^2 treba biti što bliže 1
- Pored R^2 posmatra se i vizualno poklapanje krivulja
- Najčešće se koriste polinomi 2., 3. ili 4. reda
- Treba probati više različitih funkcija dok se ne dobije što bolji R^2

ASCII tekst datoteke

- Često softver za prikupljanje podataka daje rezultate kao:
- TXT – ASCII datoteka
- TXT – Tab Delimited
- CSV – Comma Separated Values
- Za obradu podataka pomoću MS Excela potrebno je ispravno importovati podatke

ASCII tekst datoteke



ASCII tekst datoteke

■ Data, Text to Columns...

