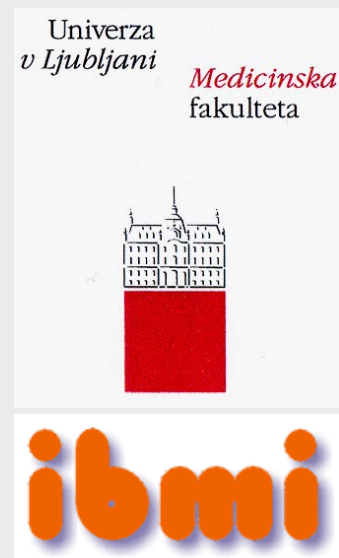


SKLADNOST – MERE IN PRIKAZ

izr. prof. dr. **Gaj Vidmar**



- Najpreprostejša mera: **delež (%) skladnosti**
 - zelo slabe statistične lastnosti
- Prve statistike za dvojiške podatke
 - Bennet *et al.* S (1954), ponovno izumljen kot
 - G (Holley & Guilford, 1964)
 - R.E. (random error, Maxwell, 1970)
 - C (Janson & Vegelius, 1979)
 - κ_n (Brennan & Prediger, 1981)
 - I_r (intercoder reliability, Perreault & Leigh, 1989)
- Scottov Π (1955) \rightarrow Cohenov κ (1960)
 - pričakovane frekvence ustrezajo **slučajni skladnosti**
 - za njihov izračun Scottov Π predpostavlja robno homogenost, Cohenov κ pa ne
 - Scottov Π $e_{ij} = N [((p_{i.} + p_{.j})/2)^2]$
 - Cohenov κ $e_{ij} = N [p_{i.} \times p_{.j}]$

■ Definicija

$$\kappa = \frac{\text{opaženi delež skladnosti} - \text{pričakovani delež skladnosti}}{(1 - \text{pričakovani delež skladnosti})}$$

$$\Leftrightarrow \kappa = \frac{o-p}{N-p}$$

- $o = \Sigma$ opaženih frekvenc na diagonali, ki predstavlja skladnost
- $p = \Sigma$ frekvenc na tej diagonali, pričakovanih po naključju
- $N =$ velikost vzorca (št. ocenjevanih/razvrščanih oseb/predmetov)

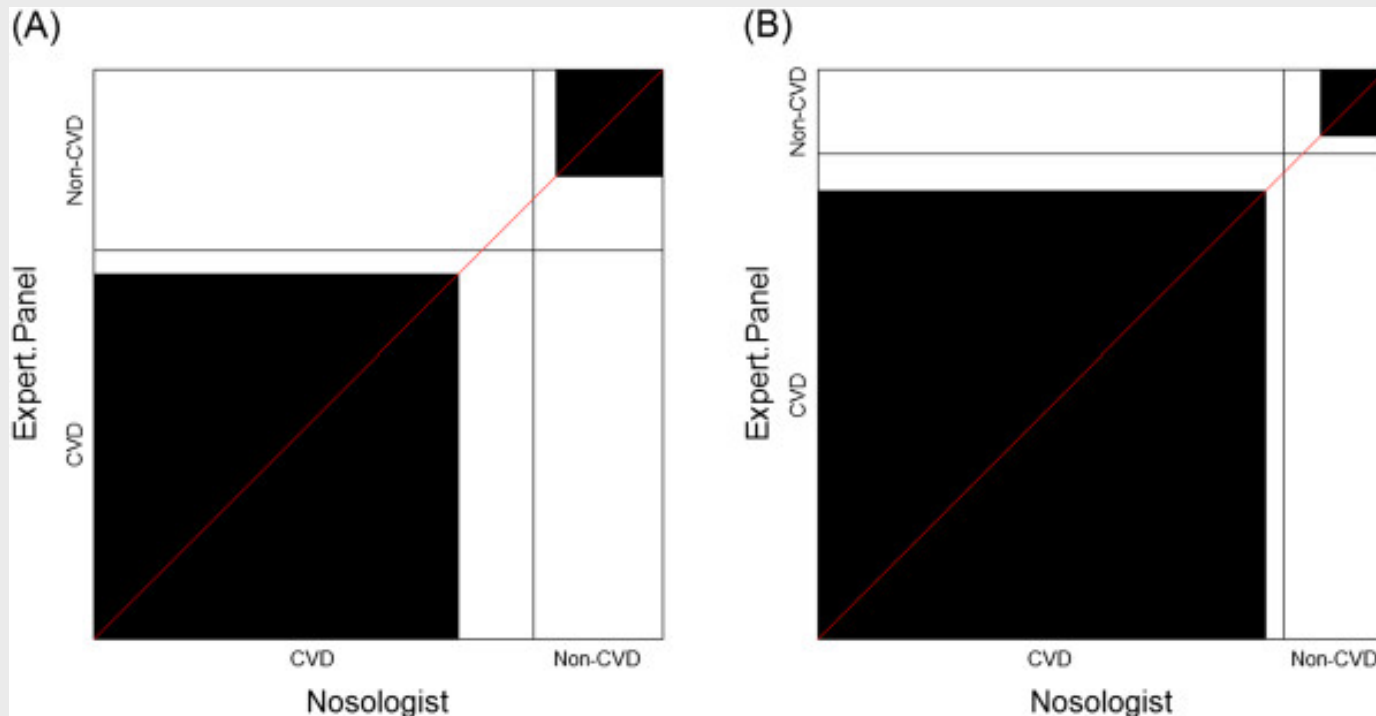
■ Povezava z McNemarovim testom

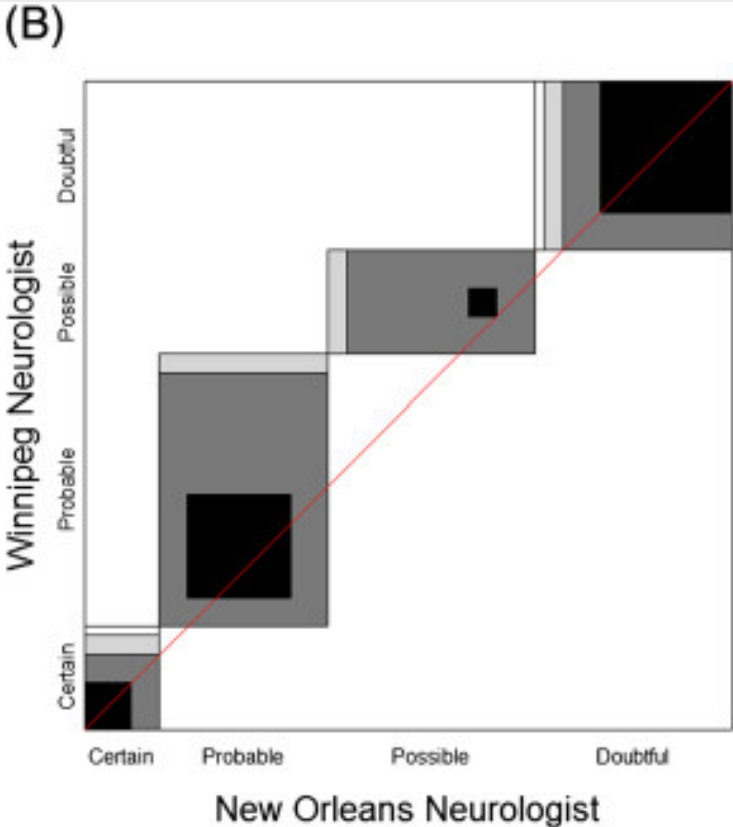
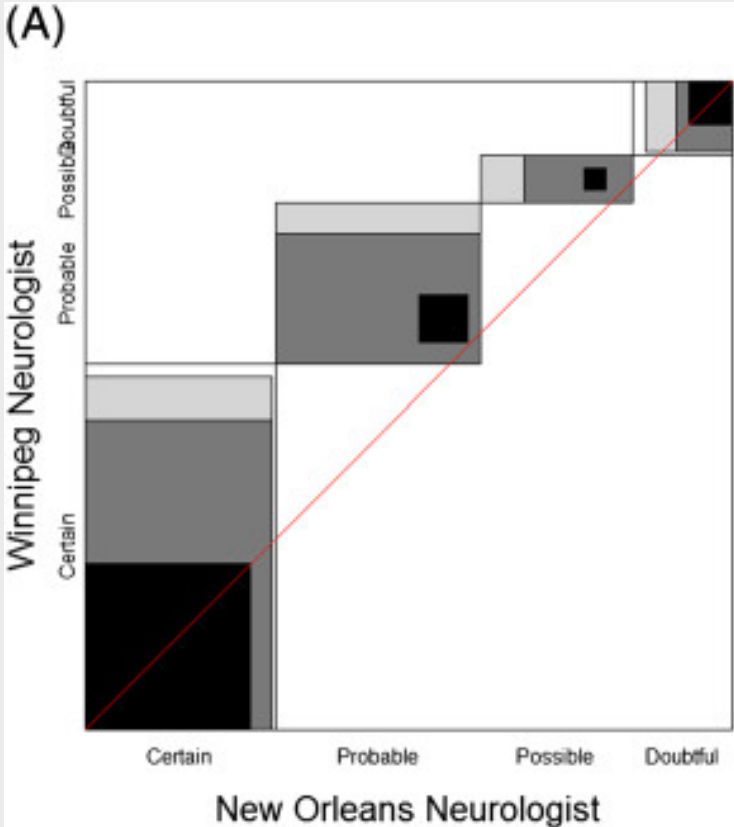
- ocenjevanji neskladni $\Leftrightarrow p_{\text{McNemar}} \rightarrow 0 \Leftrightarrow \kappa \rightarrow 0$
- ocenjevanji skladni $\Leftrightarrow p_{\text{McNemar}} \rightarrow 1 \Leftrightarrow \kappa \rightarrow 1$

- Fleissov κ
 - posplošitev Scottovega Π na več ocenjevalcev
- Razširjena κ
 - več ocenjevalcev
- Urejenostni podatki: uteženi κ
 - več kategorij
 - poseben primer je ICC

- **AC1** za opisne podatke
 - posplošljiv na več ocenjevalcev
- **AC2** za urejenostne podatke
 - posplošljiv na številske podatke
- [Teorija, programje, članki, knjiga](#)

- Za dvojiške in urejenostne podatke: [Bangdiwalov grafikon](#)
 - dva ocenjevalca oz. razvrščanji
 - nadgrajen mozaični grafikon
 - polja glede na robne frekvence, črni kvadrati skladnost
 - pri urejenostnih podatkih sivine za bližnje kategorije (delno skladnost)



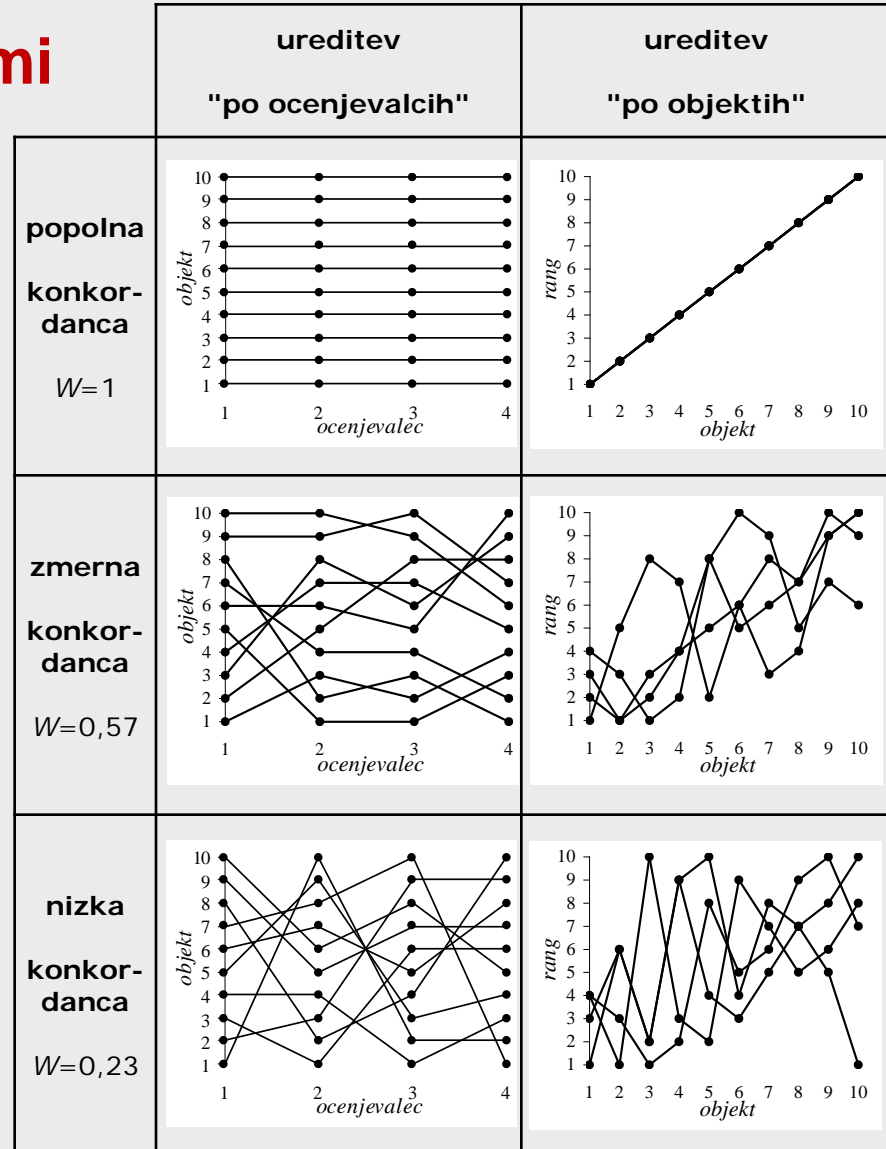


- **konkordanca = skladnost m rangiranj k objektov**
- Kendallov W in sorodne mere
 - $W \propto$ povprečni Spearmanov ρ za vse dvojice rangiranj
 - Ehrenberg: povprečje vseh Kendallovih τ
 - povprečna korelacija med ocenjevalci in kriterijem
 - testna statistika (L) in koeficient ($\mathcal{W}[-1, 1]$) za primerjavo konkordance med skupinama
- Legendreov pristop
 - naknadne primerjave dvojic ocenjevalcev (Kendallov τ , Holmov popravek vrednosti p)
 - združevanje ocenjevalcev v skupine (clustering)
- metoda Svenssonove (rehabilitacija)

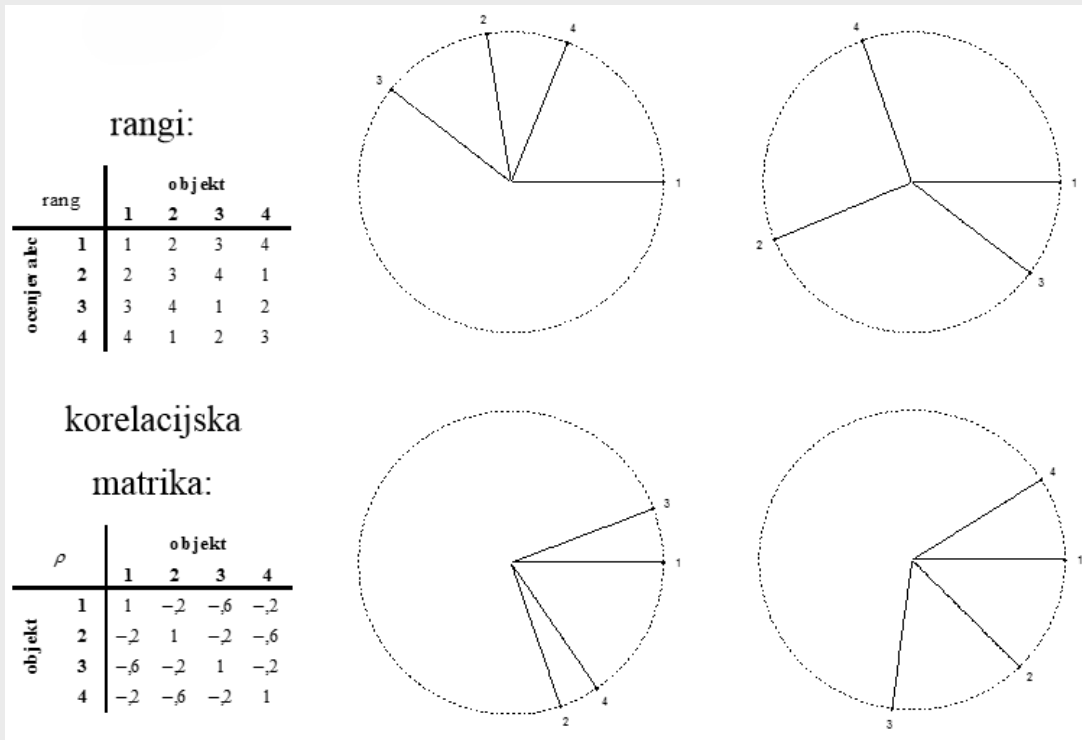
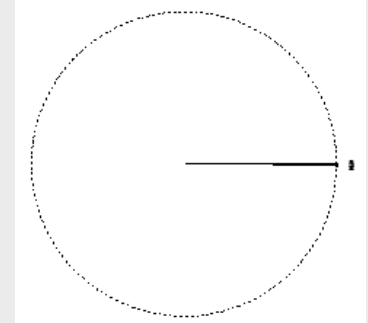
- naraščajoč pomen analize urejenostnih podatkov
- številne raziskovalne hipoteze so v resnici urejenostne narave (Cliff, *Ordinal Methods for Behavioral Data Analysis*, 1996)
- razcvet prikaza podatkov; vizualizacije konkordance še ni bilo
- prikaz dopolnjuje primerjavo konkordance med skupinami (Schucany & co-workers, 1973-84; Vidmar & Černigoj, 2004; Legendre, 2005)
- [Vidmar & Rode, Visualizing concordance, *Comput Stat*, 2007](#)

■ Prikaz z vzporednimi osmi

- osi ocenjevalci
- osi objekti

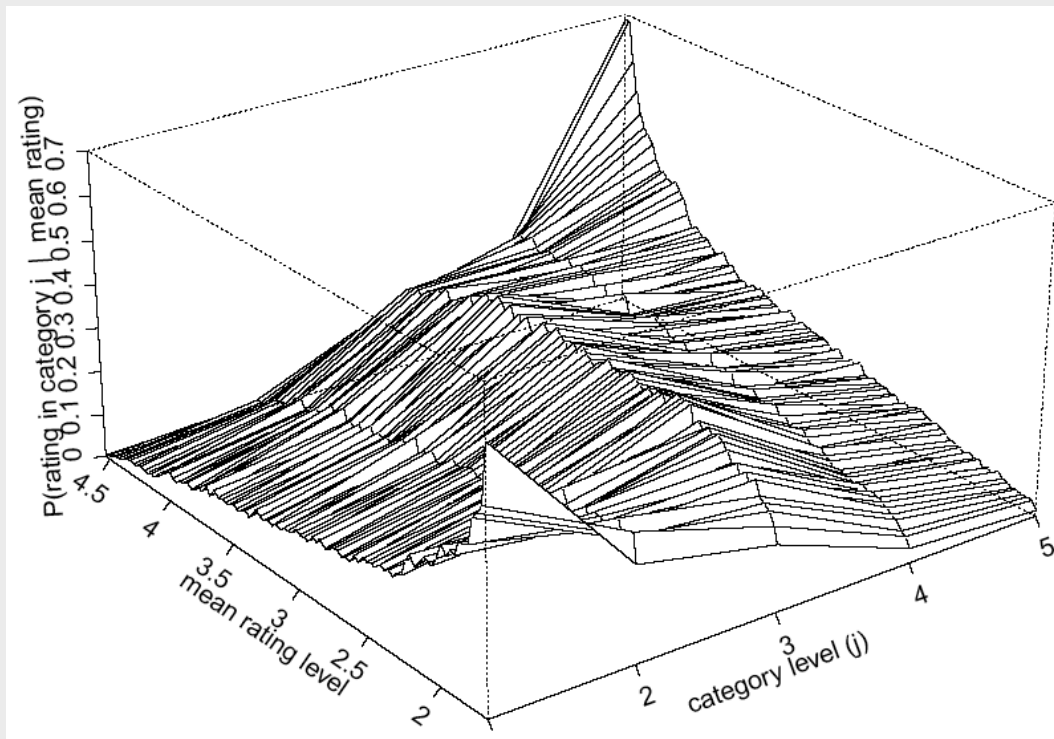


- **Trossetov prikaz korelacijskih matrik** (za Spearmanov ρ)
 - ustrezen za popolno konkordanco ($W=1$)
 - pri majhnih vrednostih k in m za nizko konkordanco algoritem (ki je soroden MDS) ne konvergira
 - za ničelno konkordanco ($W=0$) možne različne predstavitve



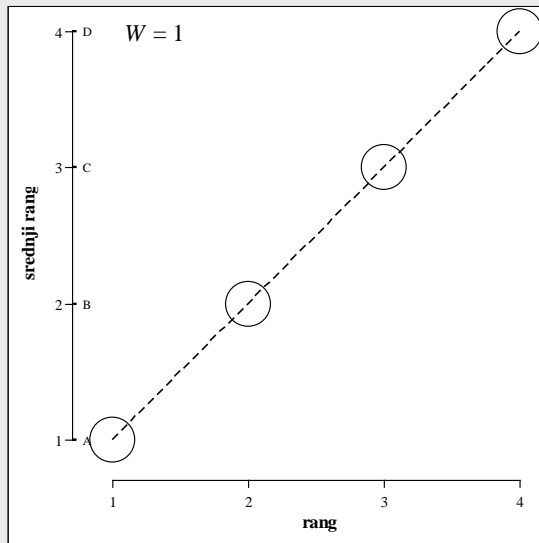
■ Nelson in Peppe

- razpršenost ocen ocenjevalcev, ki uporabljajo urejene kategorije
- zaradi soodvisnosti med povprečjem in varianco pri urejenostnih podatkih da moramo porazdelitev ocen opazovati ločeno glede na povprečno oceno ocenjevalca

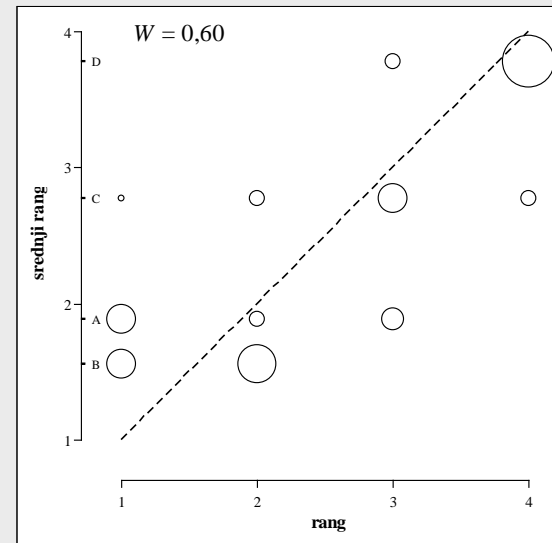


- 1. možnost: za vsak objekt narišemo vse range, ki so mu bili dodeljeni
 - **mehurčni konkordančni diagram**
 - frekvenca rangov kot funkcija povprečnega ranga
 - velikost kroga \propto št. dodeljenih rangov
 - prikaz povprečnih rangov za objekte
 - popolna konkordanca \Leftrightarrow vsi krogi na glavni diagonalni (---)
 - vprašanje merila za polmer krogov
 - alternativa je razsevni diagram z raztresenjem (jittered scatter-plot)
 - nižje razmerje podatki/črnilo
 - pri veliko ocenjevalcih nastopi prekrivanje točk

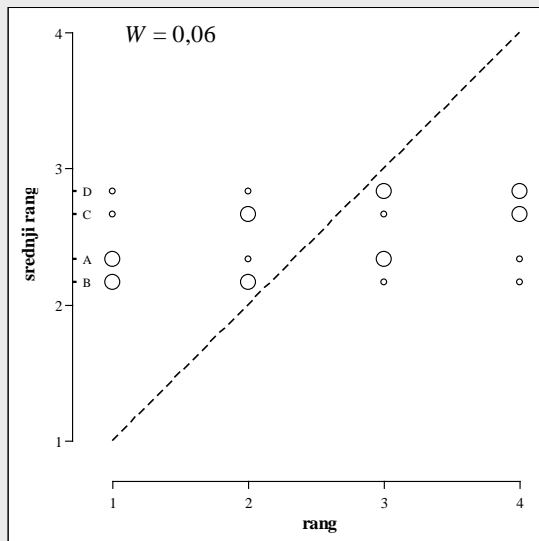
Mehurčni konkordančni diagrami



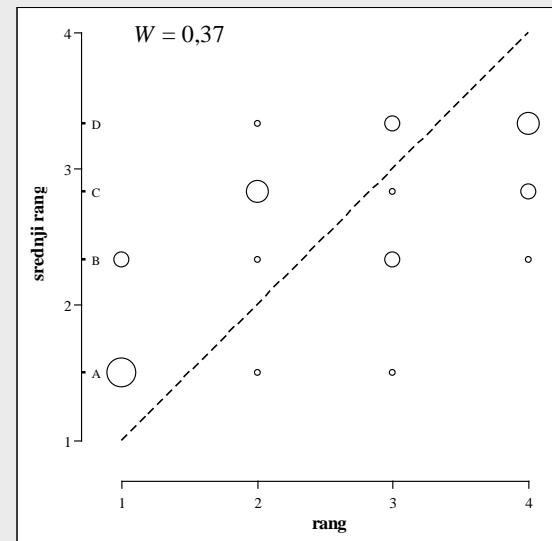
←
umetni
podatki



$k=4$
 $m=6$ ↗
 $m=9$

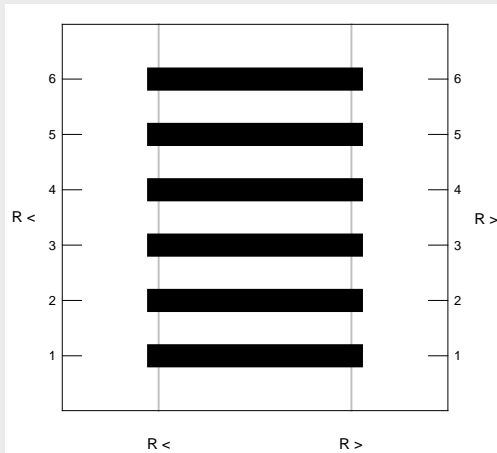


Schucany & Frawley
(1973) – US in FRA
ocenjevalci vin

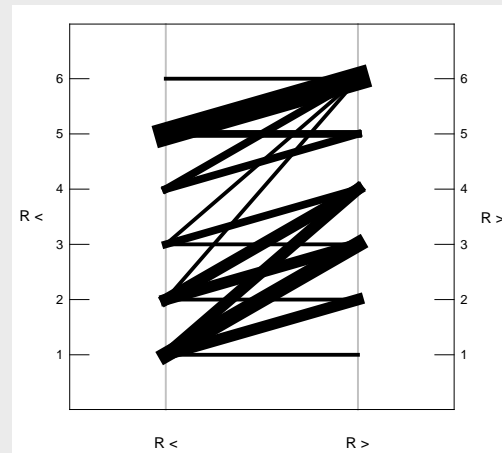


■ konkordančni diagram z vzporednima osema

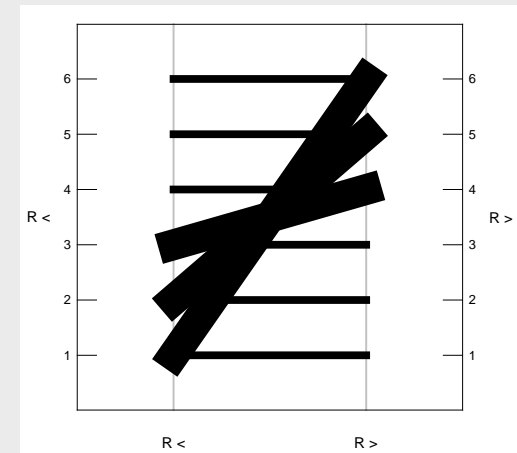
- vzporedne osi dolgo znane, popularnejše od 1990-ih (teorija Inselberg, podatkovno rudarjenje, InfoVis)
- vse dvojice rangov, v vsaki manjši ($R_{<}$) in večji rang ($R_{>}$)
- število dvojic \propto debelina črte
- $k=6$, $m=4$ (umetni podatki)



$W = 1$

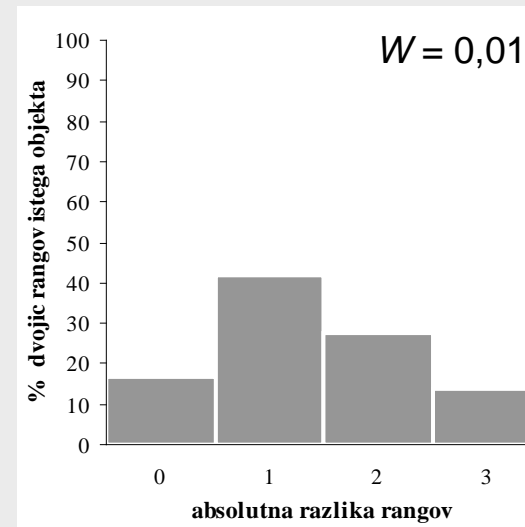
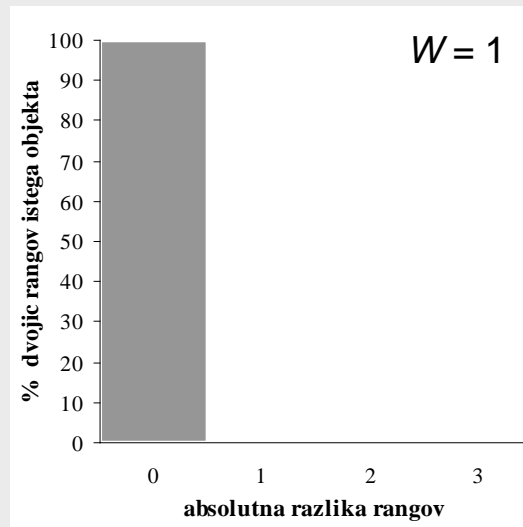


$W = 0,63$



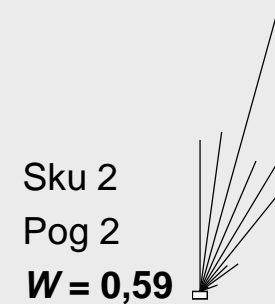
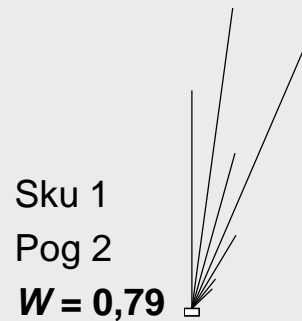
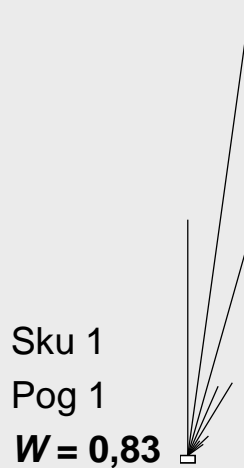
$W = 0$

- 2. možnost: rišemo razlike znotraj parov rangov
 - na njih temeljita Spearmanov ρ in Kendallov τ
 - $\forall |d_i| \in [0, k-1]$
 - $\{|d_i|\}$ ima $m(m-1)/2$ elementov
- **konkordančni stolpčni diagram**
 - absolutna razlika rangov je urejenostna spremenljivka
 - \downarrow konkordanca $\Leftrightarrow \uparrow E(|d|) \Leftrightarrow$ daljši rep porazdelitve
 - $k=4, m=9$ (umetni podatki)




■ diagram blazinice z bucikami

- polarne koordinate
 - razlika 0∞ navpična črta
 - kot v desno = $90^\circ \{1 - [|d| / (k - 1)]\}$
 - razlike ni v podatkih \Rightarrow daljice ne rišemo
 - dolžina daljice \propto št. razlik
- v izhodišču “blazinica” (širina=2, višina=1)
- $k=6$, $m=4$ (Vidmar & Černigoj, 2004, socialne norme)



- mehurčni konkordančni diagrami
 - medicina: študija soglasja med diagnostičnimi postopki

sciendo Radiology and Oncology | Ljubljana | Slovenia | www.radioloncol.com 

research article

Diagnostic accuracy of haemophilia early arthropathy detection with ultrasound (HEAD-US): a comparative magnetic resonance imaging (MRI) study

Domen Plut^{1,2}, Barbara Faganel Kotnik^{2,3}, Irena Preložnik Zupan^{2,4}, Damjana Ključevsek^{2,3}, Gaj Vidmar^{2,5}, Ziga Snoj^{1,2}, Carlo Martinoli⁶, Vlada Salapura^{1,2}

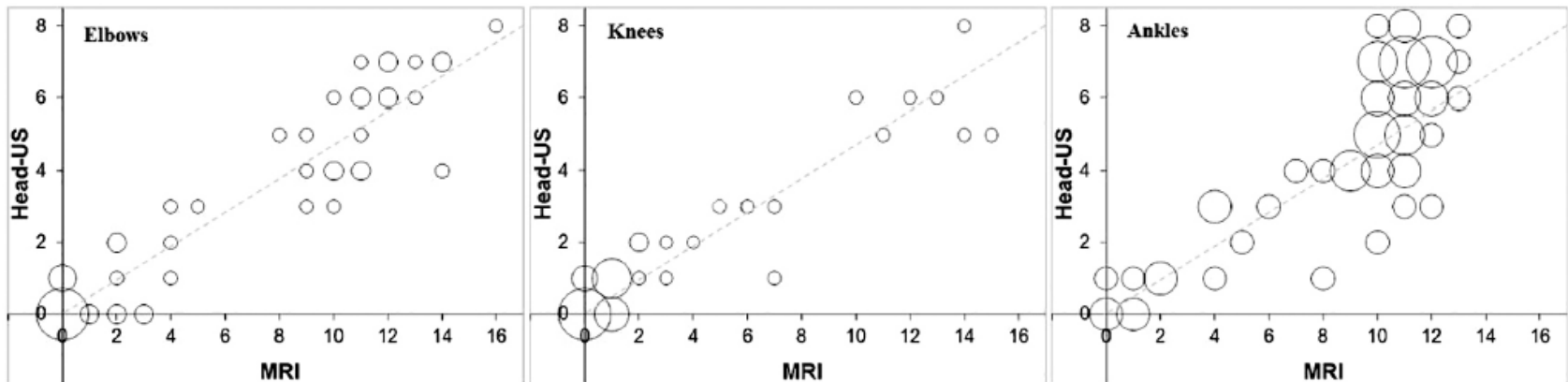


FIGURE 1. Concordance bubble-plot for depicting agreement between HEAD-US and MRI score for all three joints. The circles are centered at the observed combinations of the HEAD-US and MRI scores; their size is proportional to the number of the patients with a given combination. Dashed line represents a perfect agreement.

- intervalni podatki
- **koeficient intraklasne korelacije (ICC)**
 - ICC gre proti največji možni vrednosti 1, ko ni variabilnosti med ocenjevalci, t.j. ko vsakega ocenjevanca vsi ocenjevalci ocenijo enako, kar pomeni, da je vsa variabilnost zgolj posledica razlik med ocenjevanci
 - Če so podatki urejeni v matriko, kjer vrstice predstavljajo ocenjevance, stolpci pa ocene posameznih ocenjevalcev, bo ICC torej višji, bolj ko si bodo ocene v dani vrstici podobne
 - splošna definicija ICC =

pojasnjena varianca (t.j. varianca, ki jo pojasnjujejo dejanske razlike v merjeni lastnosti)
skupna varianca (t.j. pojasnjena varianca + varianca med ocenjevalci + variance napake)

- So ocenjevalci vsi, ki nas zanimajo, ali slučajen vzorec iz populacije potencialnih ocenjevalcev?
- So ocenjevanci vsi možni ali slučajni vzorec?
- Ali bo v praksi en ocenjevalec ali bomo uporabili povprečje več ocenjevalcev?
- 6 osnovnih oblik
 - ICC(1,1) enosmerni model za posamezno meritev
 - ICC(1,k) enosmerni model za povprečje meritev
 - ICC(2,1) dvosmerni slučajni model za posamezno meritev
 - ICC(2,k) dvosmerni slučajni model za povprečje meritev
 - ICC(3,1) dvosmerni mešani model za posamezno meritev
 - ICC(3,k) dvosmerni mešani model za povprečje meritev

- Enosmerni slučajni model (*one-way random effects model*)
 - slučajni vzorec ocenjevalcev, vsi ocenjevanci
 - uporaben tudi za ponovljena ocenjevanja
- Dvosmerni slučajni model (*two-way random effects model*)
 - slučajni vzorec ocenjevalcev in ocenjevancev
 - najbolj posplošljiva ocena ICC
- Dvosmerni mešani model (*two-way mixed model*)
 - vsi ocenjevalci, slučajni vzorec ocenjevancev
 - ocena ICC ni posplošljiva na druge ocenjevalce

- Predvidena prihodnja izvedba ocenjevanj oz. meritev:
 - posamezna meritev (*single measure*)
 - povprečje meritev (*average measure*)
- Dodatno pri vseh štirih dvosmernih modelih ločimo, ali želimo ocenjevati
 - absolutno strinjanje (ko nam pomeni pristranost, t.j. različno strog kriterij med ocenjevalci, vir nezanesljivosti) ali
 - konsistentnost (ko variance med ocenjevalci, ki izhaja iz pristranosti, ne štejemo v varianco napake);
 - tako dobimo skupaj 10 oblik ICC
- Cronbachov α je posebna oblika ICC

■ Skladnost med metodami

■ postopek Blanda in Altmana

- navaden
- za ponovljene meritve

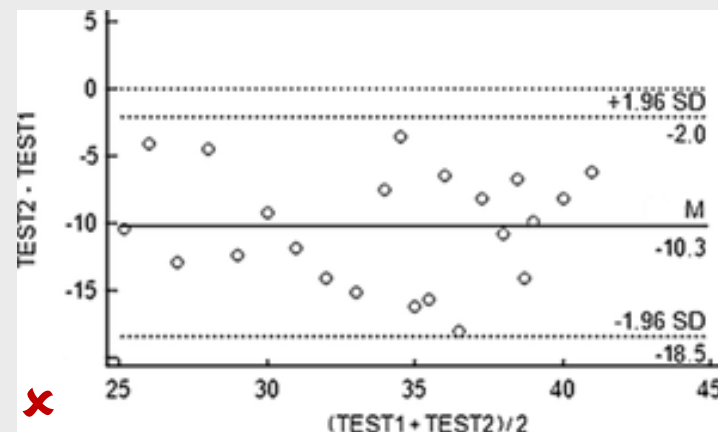
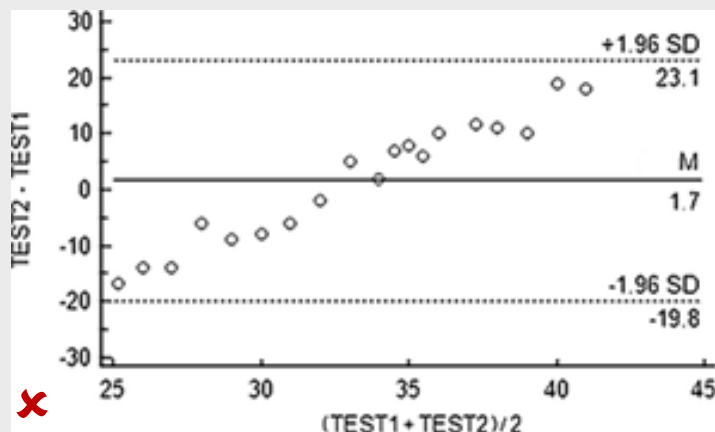
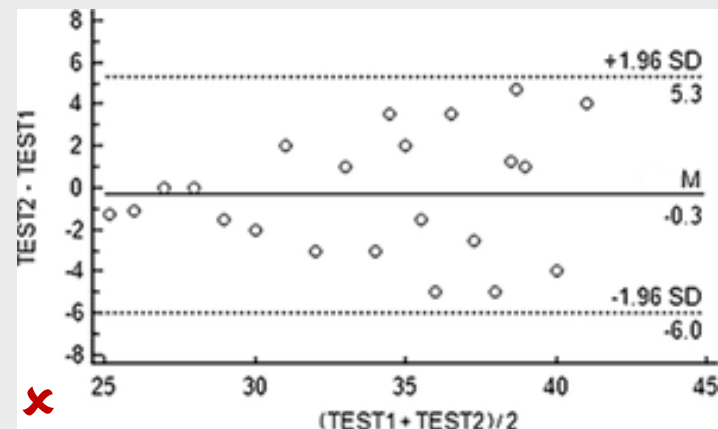
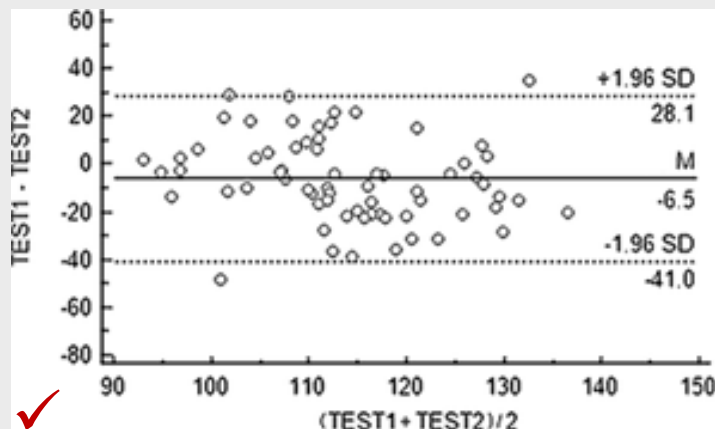
■ regresija

- navadna skozi izhodišče (primerjamo z diagonalo)
- Demingova
- Passing in Bablock (neparametrična)
- različna poimenovanja

- *model I* = regresija po metodi najmanjših kvadratov (*ordinary least squares*, OLS; napake so le pri odvisni spremenljivki); *model II* = napaka tudi pri odvisni spremenljivki
- glavna pristopa za model II sta metoda glavne osi (*major axis*, MA) in metoda zmanjšane glavne osi (*reduced major axis*, RMA)
- naklon regresijske premice po metodi RMA je geometrijska sredina naklonov b_{yx} in b_{xy} , zato se imenuje tudi *geometric mean regression* (GM); sinonimi za RMA so tudi metoda najmanjših zmnožkov (*least products*), najmanjših pravokotnikov (*least rectangles*) in standardizirana regresija glavnih komponent (*standardised principal component regression*, SPCR)
- RMA je poseben primer splošnega Demingovega modela (najpogosteje uporabljan; v fiziki je v zvezi s to metodo najpogosteje citiran avtor York)
- dodatna možnost je povprečje naklonov b_{yx} in b_{xy} (*arithmetical mean regression*, AMR)
- vse štiri regresijske premice (OLS, MA, RMA, AMR) gredo skozi točko $\{M_x, M_y\}$

■ Linov koeficient skladnostne korelacije (CCC)

- razlika v odvisnosti od povprečja (*mean-difference plot*)
 - meje skladnosti (*limits of agreement*): $M \pm 1,96$ SD razlik
 - [eden najbolj citiranih člankov nasploh](#)
 - sodobna dopolnitev z regresijo ([Francq & Govearts](#))

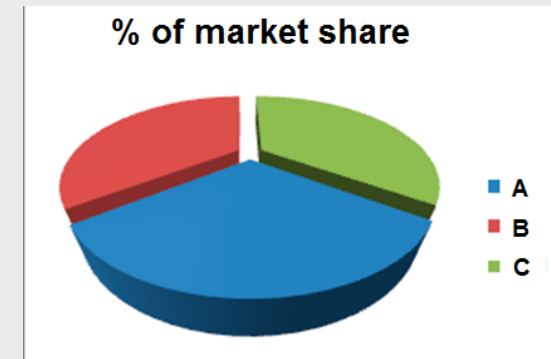
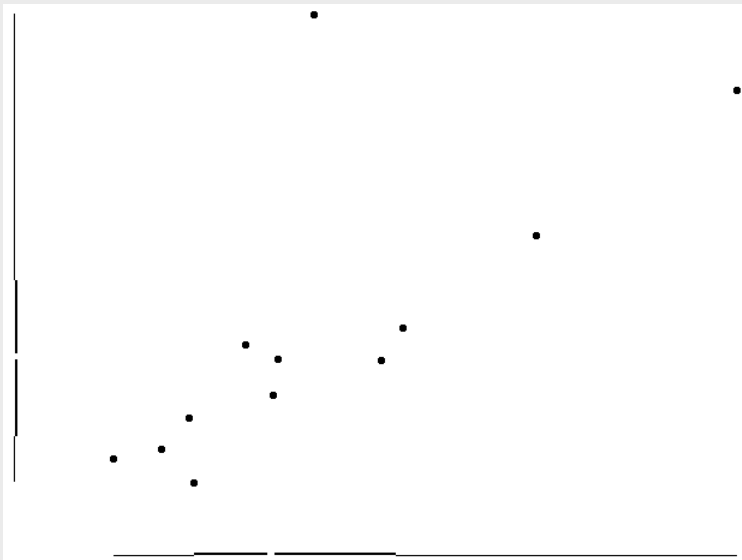


- Robustne in posplošene oblike CCC
 - neparametrična, vključuje krnjenje
 - robustno ocenjena razdalja (Huber, Winsor)
 - posplošena verzija za ponovljena merjenja (King, tudi za binarne podatke)
- **Krippendorfov α**
 - [Hayes & Krippendorf](#)

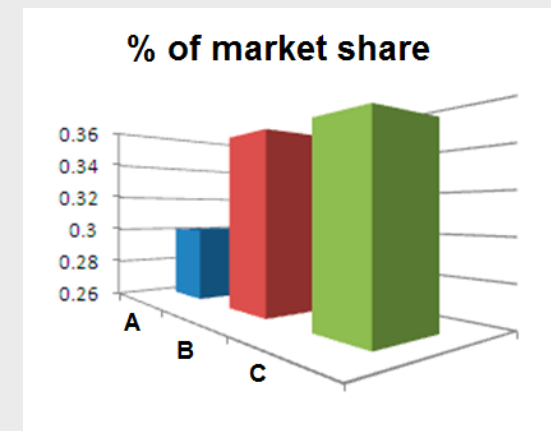
- [Agreestat](#) (Excel, spletni, paket za R, SAS)
- [SPSS](#)
 - vsi ICC, Cronbachov α idr. mere zanesljivosti, Kendallov W in τ , Cohenov κ
 - številni makroji ([M.García-Granero](#))
 - makro za Krippendorfov α ([A.F.Hayes](#))
- R
 - paket [irr](#)
 - paket [MethComp](#)
- [MedCalc](#)
 - κ , ICC, Linov CCC
 - B&A in [diagram gore](#)
 - regresiji za primerjavo metod
- Minitab
 - Cohenov in Fleissov κ , Kendallov W in τ ([Attribute Agreement Analysis](#))
 - pristopi QC/SPC ([Gage R&R, MSA](#))
- [SAS](#), [Stata](#), [NCSS](#), [StatsDirect](#), [GenStat](#)
- Dodatki za Excel: [XLSTAT](#), [Analyse-it](#), [Real Statistics Using Excel](#)

■ Pionir: Edward Tufte


- grafična odličnost vs. **risbosmetje (chartjunk)**
 - **faktor lažnjivosti (lie factor)**
= velikost učinka na grafikonu / v podatkih
! naj bo čim bližje 1
 - **razmerje podatki-črnilo (data-ink ratio)**
= črnilo za podatke / skupno črnilo na grafikonu
! naj bo čim večje
 - **iskrne črte (sparklines)**
= preprosti, a sporočilni miniaturni grafikoni (v besedilu)
- } jasno in učinkovito sporočiti kompleksne ideje



FL >> 1



MSFT 70-week: 26.45  28.35 [18.95|31.63]

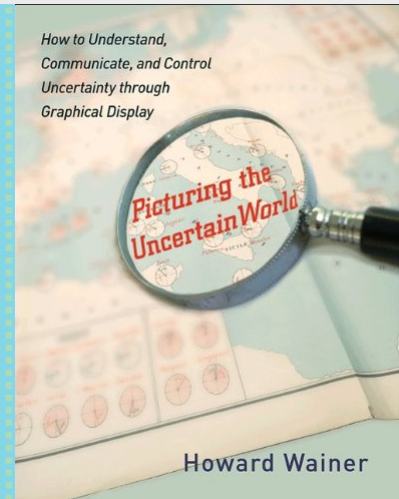
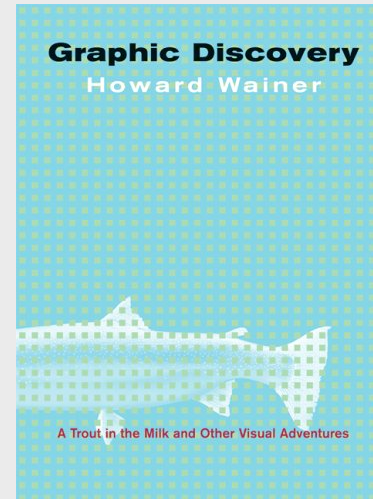
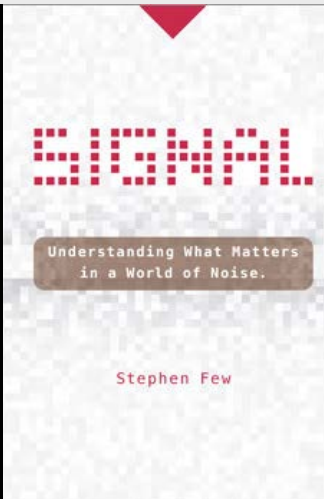
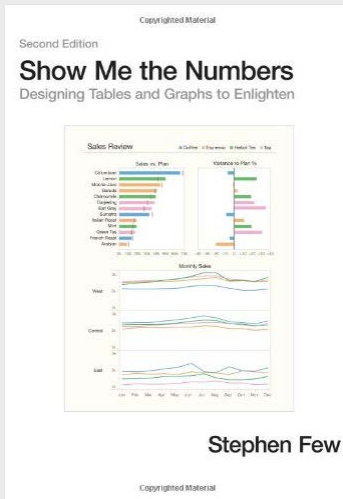
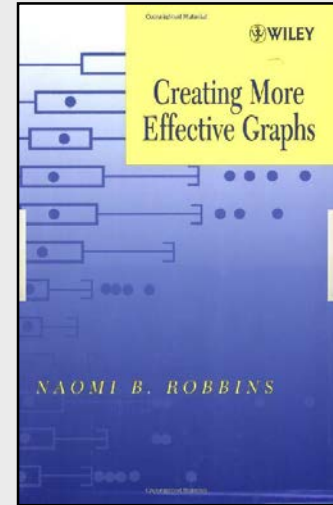
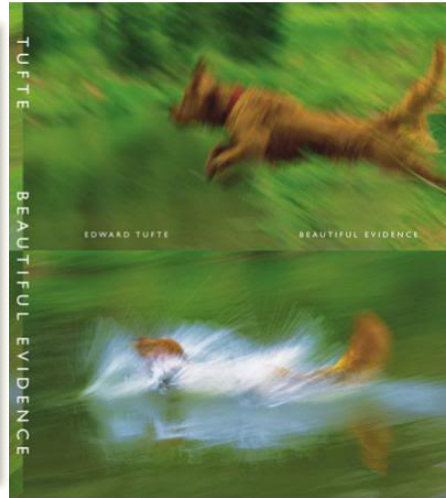
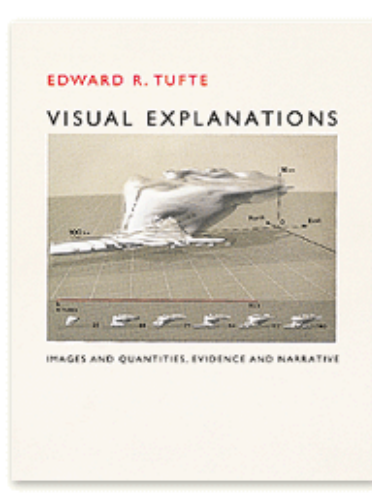
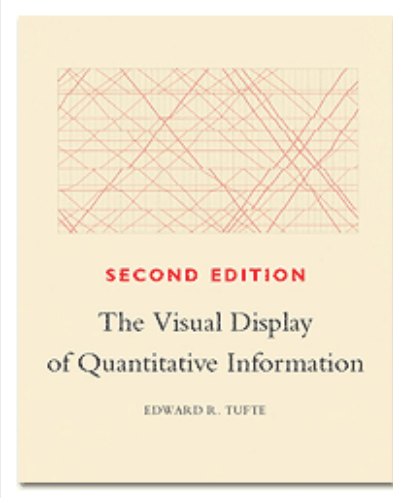
Overall Rating (1-5):  3.56

Win-Loss-Tie: 7-5-2 

Osnovne knjige o prikazu podatkov

DODATEK

28 / 26



Sodobna teorija in praksa prikaza pod.

