

Analiza podataka

Osnovni pojmovi, *Excel* - Program za tabelarne proračune, *Python*, *Pandas*

Uvod u informacijski i finansijski inženjering

1

Koraci u analizi podataka

- Analiza podatka
 - je proces koji obuhvata uvid u podatke, prečišćavanje, transformaciju i modelovanje podataka radi otkrivanja korisnih informacija a sa ciljem pružanja podrške poslovanju
 - proces analize obuhvata sledeće korake:
 1. Definisanje istraživačkog cilja
 2. Prikupljanje podataka
 3. Priprema podataka
 4. Istraživanje podataka
 5. Kreiranje modela
 6. Prezentovanje rezultata

2

2

Koraci u analizi podataka

- 1. Definisanje istraživačkog cilja
 - šta istražujemo
 - koja je korist tog istraživanja za kompaniju
 - koji podaci su potrebni
 - šta je konačan cilj istraživanja
 - koji su rokovi

3

3

Koraci u analizi podataka

- 2. Prikupljanje podataka
 - kompanije imaju svoje podatke, koji ne moraju biti iz jednog izvora
 - postoje i javni/besplatni skupovi podataka koji se mogu kombinovati
 - <https://www.kaggle.com/datasets>
 - <https://catalog.data.gov/dataset>
 - <https://data.europa.eu/euodp/data/dataset>
 - <https://datacatalog.worldbank.org/dataset/>
 - <https://datasetsearch.research.google.com/>
 - postoji problem **privatnosti** podataka
 - potrebno je odvojiti vreme za sređivanje podataka
 - u ovom koraku nas zanima da li imamo potrebne podatke i da li su oni u odgovarajućem formatu

4

4

Koraci u analizi podataka

- 3. Priprema podataka
 - čišćenje podataka
 - greške nastale pri unosu
 - npr. višak razmaka na početku ili kraju vrednosti (strip() funkcija u Python-u)
 - velika i mala slova
 - nerealne vrednosti
 - autlajeri (outliers)
 - vrednosti koje nedostaju
 - odstupanje od predviđenih vrednosti
 - različite jedinice za merenja
 - različiti načini agregacije

5

5

Koraci u analizi podataka

- 3. Priprema podataka
 - kombinovanje različitih skupova podataka
 - postoje dva načina ako govorimo o tabelarnim podacima
 - spajanje skupova na osnovu zajedničkog obeležja
 - dodavanje jednog skupa na kraj drugog
 - upotreba materijalizovanih pogleda radije nego fizičko spajanje
 - dodavanje novih obeležja na osnovu agregiranih vrednosti

6

6

Koraci u analizi podataka

- 3. Priprema podataka
 - transformisanje podataka
 - izbacivanje osobina koje imaju visoku korelaciju
 - primena logaritama na nezavisne promenljive
 - pretvaranje vrednosti u "dummy" vrednosti
 - ...

7

7

Koraci u analizi podataka

- 4. Istraživanje podataka
 - cilj je bolje razumeti zavisnost između podataka
 - grafički prikaz podataka može biti od velike koristi

8

8

Koraci u analizi podataka

- 5. Kreiranje modela
 - cilj kreiranja modela je da se obezbedi predikcija, da se klasifikuju novi elementi na osnovu postojećih
 - u ovom koraku primenjuju se tehnike mašinskog učenja, data mininga, statistike, ...
 - kreiranje modela sastoji se od:
 - izbora tehnike za modelovanje i promenljivih koje učestvuju u izgradnji modela
 - implementacije i izvršavanja modela
 - analiziranja dobijenih rezultata

9

9

Koraci u analizi podataka

- 6. Presentovanje rezultata
 - vizualizacija rezultata istraživanja
 - automatizacija celog postupka kako bi neko drugi mogao da ponovi postupak

10

10

Tipovi podataka

- **Raznolikost podataka**
 - izuzetno je bitna jer utiče na **način na koji se obrađuju podaci**
 - tipovi podataka prema strukturi
 - **strukturirani podaci**
 - **nestrukturirani podaci**
 - **polustrukturirani podaci**

11

11

Tipovi podataka

- Tipovi podataka prema strukturi
 - **strukturirani podaci**
 - postoji **šema** koja definiše format podataka
 - striktno zadovoljavaju predefinisani format
 - u praksi, očekuje se da takva šema bude razrađena do potrebnog stepena detaljnosti
 - uobičajeno **pogodni za analizu**
 - primer
 - podaci u relacionoj bazi podataka
 - sve torke imaju isti format

12

12

Tipovi podataka

Structured

Conforms to a
schema



Order	CustID	Month	Item	Color	Price
101	20051	Dec	Pen	Red	2.99
102	20045	Mar	Pencil	Blue Yellow Red	3.99
103	29584	May	Eraser	Blue	1.25
104	29584	May	Pen	White	2.25
105	29584	May	Pencil	Blue Yellow Red	2.99
106	27485	Jan	Eraser	Blue Yellow	2.75
107	29574	Jan	Marker	Green	1.75
108	24447	Feb	Marker	Yellow Blue	7.25
109	26466	Jul	Pen	Black Red	5.25
110	27467	Jun	Pencil	Black	2.95

izvor: *Fundamentals of Big Data, Databricks*. <https://academy.databricks.com/>

13

13

Tipovi podataka

- Tipovi podataka prema strukturi
 - **nestrukturirani podaci**
 - (A) može postojati šema koja definiše format podataka
 - tipovi podataka su, međutim, često "preširoki"
 - nedovoljno detaljni sa stanovišta semantike podataka
 - (B) ne postoji ni šema koja definiše format podataka
 - ne postoji nikakav ugrađeni format kojeg podaci prate
 - moguće je da postoji nekakav eksterni format koji opisuje semantiku podataka, ali on ne pripada strukturi
 - **najrasprostranjeniji** tip podataka
 - po IDC, 90% svih podataka su nestrukturirani
 - primeri
 - objave sa socijalnih mreža ili email poruke
 - tekstualni ili multimedijalni dokumenti
 - podaci su nestrukturirani sa stanovišta sistema za njihovo skladištenje

14

14

Tipovi podataka

Unstructured

Does not fit neatly
Into a schema



izvor: *Fundamentals of Big Data, Databricks*. <https://academy.databricks.com/>

15

15

Tipovi podataka

- Tipovi podataka prema strukturi
 - **polustrukturirani podaci**
 - imaju definisanu strukturu ali postoje i odstupanja
 - atributi mogu da postoje
 - u svim entitetima
 - samo u nekim entitetima
 - prikazuju se, često, putem grafova i stabala
 - moguće je definisati šemu
 - koja specificira moguće elemente
 - koji učestvuju u strukturi
 - mogu ali ne moraju svi da postoje
 - primer
 - HTML kôd

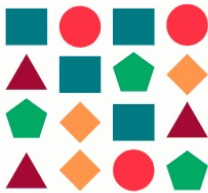
16

16

Tipovi podataka

Semi-structured

Some level of organization



```

▼<div class="new-main-menu">
  ▼<div class="header-desktop-block">
    ▼<div class="container new-menu">
      ▶<a class="main-logo" rel="home" href="https://databricks.com/" title="Databricks">...</a>
      ▼<div id="new-m" class="menu-bar">
        ▼<div id="mega-menu-wrap-headerNew" class="mega-menu-wrap">
          ▶<div class="mega-menu-toggle">...</div>
          ▼<ul id="mega-menu-headerNew" class="mega-menu max-mega-menu mega-menu-horizontal" data-event="hover_intent" data-effect="fade_up" data-effect-speed="200" data-effect-mobile="disabled" data-effect-speed-mobile="0" data-panel-width="body" data-panel-inner-width="#new-m" data-mobile-force-width="false" data-second-click="close" data-document-click="collapse" data-vertical-behaviour="standard" data-breakpoint="1199" data-unbind="true">
            ▶<li class="mega-main-bar-li mega-menu-item mega-menu-item-type-custom mega-menu-item-object-custom mega-menu-item-has-children mega-menu-megamenu mega-

```

izvor: *Fundamentals of Big Data, Databricks.* <https://academy.databricks.com/>

17

17

Microsoft Excel

18

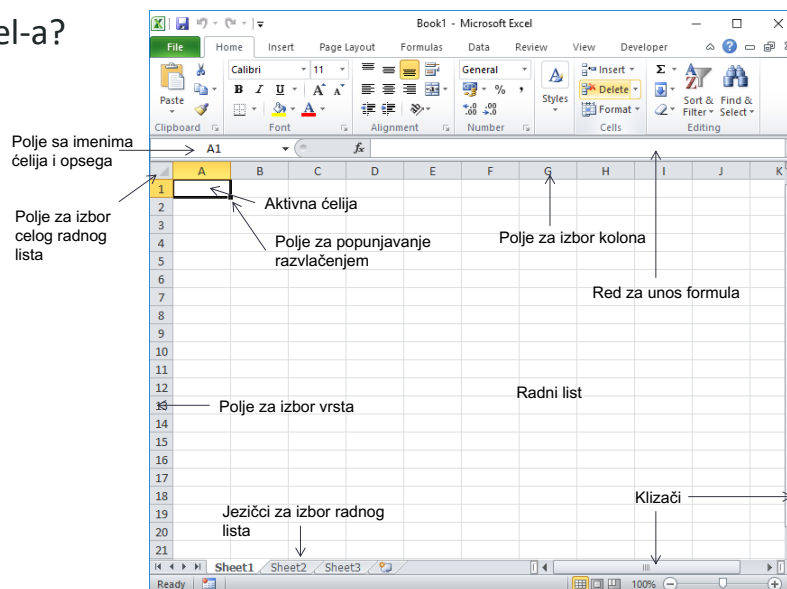
Šta je Excel?

- MS Excel je program za tabelarne proračune
 - Računanje srednje ocene učenika u odeljenju
 - Vođenje finansija preduzeća
 - Statistički proračuni
 - ...
- MS Excel poseduje veliki broj ugrađenih funkcija i omogućava
 - Manipulaciju podacima
 - Crtanje grafikona
 - Programiranje
- **Analiza strukturiranih podataka!**

19

19

Kako izgleda prozor Excel-a?



20

20

Kako izgleda prozor Excel-a?

- Dokument u Excel-u je radna sveska (eng. *workbook*)
 - datoteka u kojoj je sačuvan dokument ima ekstenziju **.xlsx**
 - u starijim verzijama **.xls**
- Sadrži jedan ili više radnih listova (eng. *worksheets*)
 - jedan radni list je uvek aktivan
 - mogu se dodavati novi listovi ili brisati postojeći

21

21

Oblici pokazivača u Excel-u

- Pokazivač miša može da ima različite oblike, što zavisi od konkretne tekuće akcije
 - A. izbor aktivne ćelije
 - B. unos uzastopnih vrednosti u grupu ćelija razvlačenjem
 - ili kod kopiranja vrednosti jedne ćelije u grupu ćelija
 - C. unos podataka u ćeliju ili u red za unos formula
 - D. odabir alata iz palete sa alatima
 - E. Promena veličine kolona ili redova



22

22

Ćelije i ćelijske reference

- Radni list može da sadrži najviše 1,048,576 redova i 16,384 kolona
- **Ćelija** – u preseku svakog reda i svake kolone
 - **koordinate ćelija** – ćelijske reference
 - Unutar jednog radnog lista
 - A1, A2, B1...
 - Između radnih listova
 - Sheet1!A1
 - moguća dodela imena skupu ćelija
 - koje se koristi **umesto** ćelijske reference

23

23

Sadržaj i izgled ćelije

- Sadržaj ćelije može biti:
 - broj
 - datum
 - tekst
 - formula
- Svaka ćelija može da se formatira
 - menjaju se atributi koji utiču na izgled ćelije

24

24

Izmena sadržaja ćelije

- Da bi sadržaj bio unet, ćelija mora biti aktivna
 - u polju sa imenom je referenca aktivne ćelije
 - označeni su nazivi kolona i redova u kojima se nalaze aktivne ćelije

	A	B	C	D	E
1					
2	123				
3					
4					

25

25

Kopiranje i premeštanje sadržaja ćelije

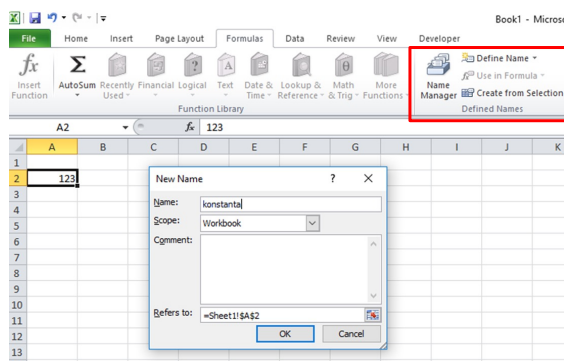
- Obeležiti željeni opseg
- Komande za rad sa klipbodom (eng. *clipboard*)
 - Copy (*CTRL + C*)
 - Cut (*CTRL + X*)
 - Paste (*CTRL + V*)
- Prevlačenje
 - pokazivač miša na jednu od ivica obeleženog opsega (menja se izgled strelice)

26

26

Definisanje imena koja se mogu koristiti u formulama

- Moguće je imenovati i ćeliju, da bi kasnije to ime moglo biti korišćeno kao parametar funkcije
 - umesto klasične reference
 - moguće je definisati doseg imena
 - radna sveska
 - radni list

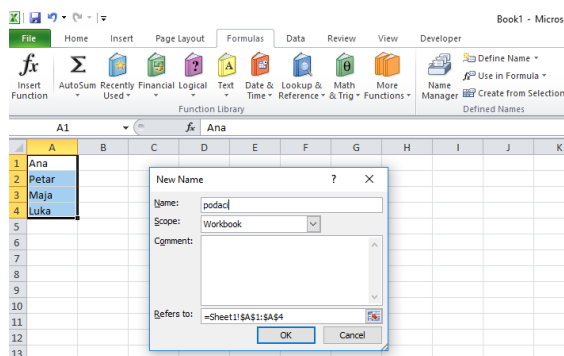


27

27

Definisanje imena koja se mogu koristiti u formulama

- Moguće je imenovati i ceo opseg ćelija, da bi kasnije mogao biti korišćen kao parametar funkcije koja radi nad opsezima



28

28

Rad sa formulama

- Običaj je da se podaci koji se mogu izračunati ne unose ručno
 - pravilo za izračunavanje – **formula**
- **Prvo se unosi znak '='**, a iza njega sledi formula
 - sastoji se od brojeva, konstanti, imena opsega, ćelijskih referenci, operatora, funkcija i zagrada, koje utiču na redosled izračunavanja

29

29

Rad sa formulama

- Kada se u ćeliju unese formula, ona se automatski izračunava i ćelija prikazuje izračunatu vrednost
 - vrednost ćelije se automatski ažurira kada se promeni neka od vrednosti koje učestvuju u formuli

	A	B	C	D	E
1	120				
2	215				
3	158				
4					
5	297				
6					

Formula bar: f_x =2*A1+A2-A3

30

30

Apsolutne i relativne ćelijske reference

- Ako se ćelija koja sadrži formulu kopira ili premešta, ćelijske reference iz formule se prilagođavaju na odgovarajući način
 - **relativne reference** – npr. A5
 - ne pamte se tačne adrese ćelija u formuli, nego njihove relativne pozicije u odnosu na ćeliju u kojoj se nalazi formula
 - ne važi kada se u formuli koristi imenovani opseg ćelija

		A6		fx		=A1+A2+A3+A4+A5	
	A	B	C	D	E	F	
1	5		1				
2	5		2				
3	3		4				
4	4		3				
5	5		1				
6	22		=C1+C2+C3+C4+C5				
7							

31

31

Apsolutne i relativne ćelijske reference

- Znakom '\$' mogu se fiksirati oznake vrsta i kolona
 - tada se fiksirana vrsta ili kolona ne menja prilikom kopiranja i premeštanja
 - **apsolutne reference** – npr. \$A\$5
 - **mešovite reference** – npr. \$A5 ili A\$5

		D3		fx		=C3*\$D\$1	
	A	B	C	D	E		
1	radni sati		din/čas	150			
2	redovno	prekovremeno	ukupno	dinara			
3	10	2	12	1800			
4	8	0	8	1200			
5	4	5	9	1350			
6							

32

32

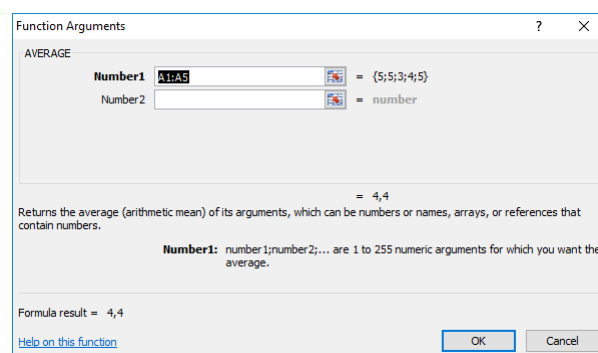
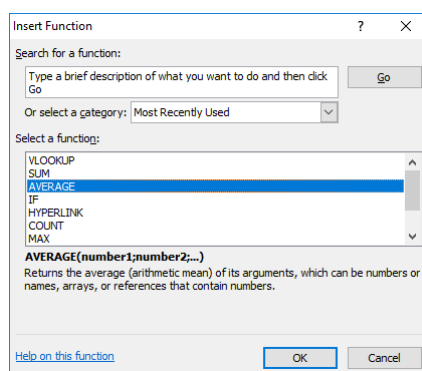
Funkcije

- Veliki broj ugrađenih funkcija
- Funkcija je određena imenom i nizom argumenata
- Unosi se
 - direktno sa tastature
 - komandom Formulas -> Insert Function
 - dugmetom sa oznakom fx, koje se nalazi u redu za formule

33

33

Funkcije

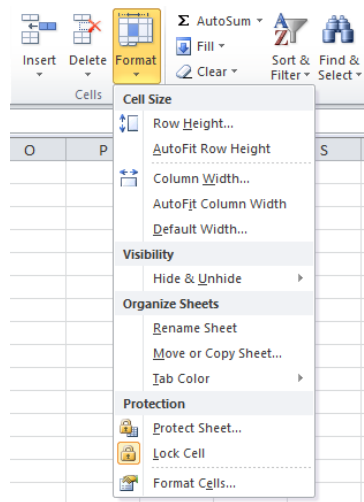


34

34

Formatiranje

- Pod formatom nekog objekta (ćelije, vrste, kolone ili radnog lista) podrazumeva se skup osobina koje određuju kako će taj objekat izgledati

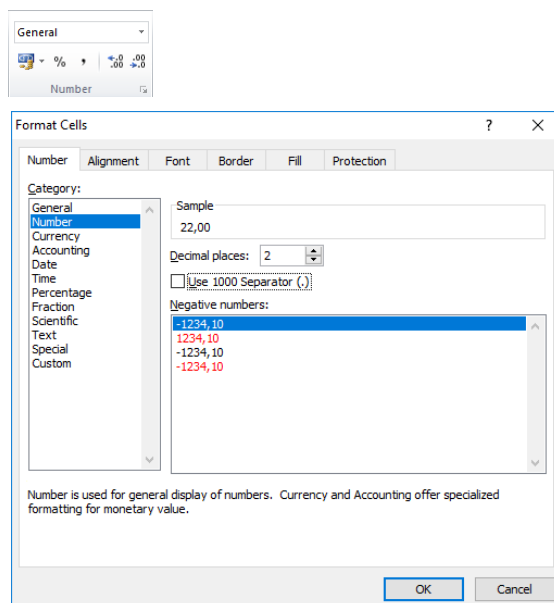


35

35

Formatiranje ćelije

- Određuje izgled ćelije i način prikazivanja podataka u njoj
 - Number* – format broja
 - Alignment* – poravnavanje
 - Font* – oblik slova
 - Border* - okviri
 - Fill* – bojenje i senčenje
 - Protection* – zaštita
- Formatiranje brojeva

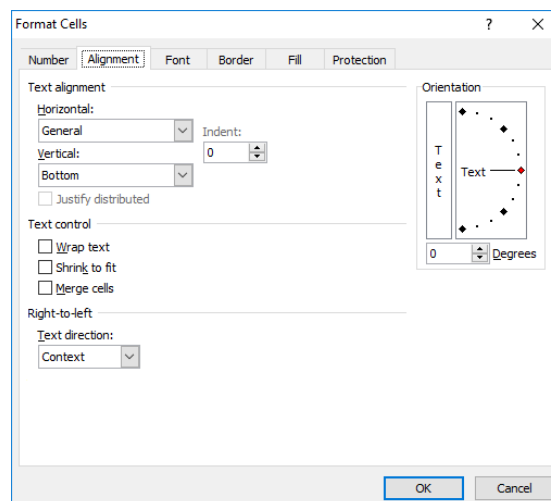


36

36

Formatiranje ćelije

- Poravnavanje ispisa u ćelijama
 - tekstualni podaci po levoj ivici
 - brojučani podaci po desnoj ivici
 - logičke vrednosti i greške centrirano



37

37

Formatiranje ćelije

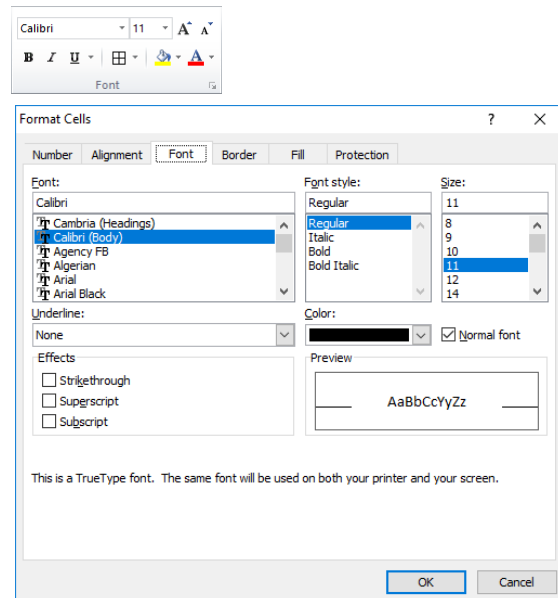
- Poravnavanje ispisa u ćelijama
 - **Wrap text**
 - povećava se visina ćelije, a širina se zadržava
 - ako je sadržaj ćelije duži od širine (biće prikazan u više redova)
 - **Shrink to fit**
 - automatsko podešavanje veličine fonta prema širini ćelije
 - **Merge and center cells**
 - spajanje više ćelija u jednu i centriranje sadržaja
 - potrebno je prvo obeležiti ćelije koje treba da se spoje, ali samo jedna može imati sadržaj

38

38

Formatiranje ćelije

- Izbor fonta, boje slova, načina ispisa

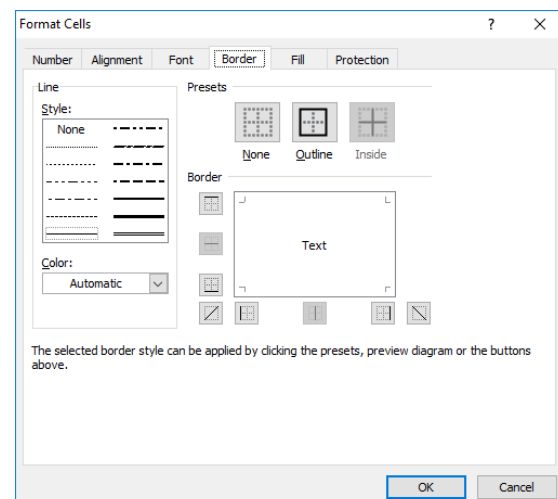


39

39

Formatiranje ćelije

- Podешavanje okvira ćelija i opsega

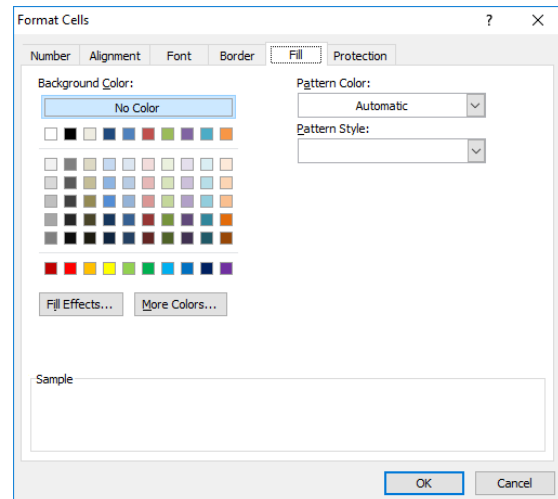


40

40

Formatiranje ćelije

- Bojenje i senčenje ćelija

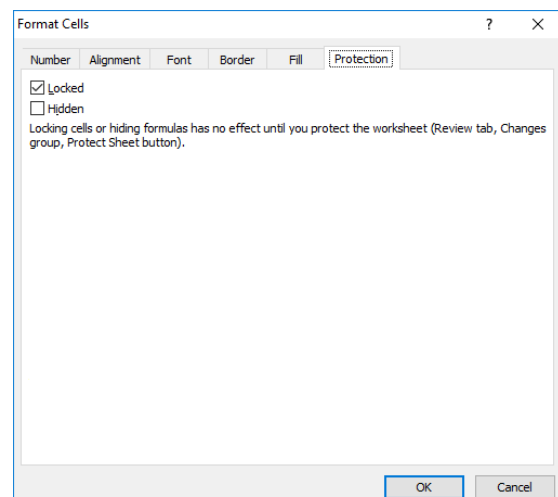


41

41

Formatiranje ćelije

- Zaštita (*Protection*) omogućava
 - zaključavanje (*Locked*) i
 - skrivanje (*Hidden*) sadržaja ćelije
- Da bi zaštita bila primenjena, potrebno je uključiti zaštitu na nivou radnog lista
 - kada je zaštita uključena, ćelije koje su zaključane ili sakrivene postaju nedostupne korisniku
 - posebno važno za formule

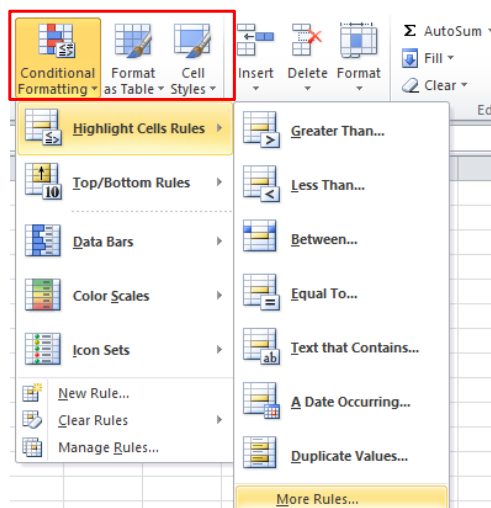


42

42

Formatiranje ćelije

- Ćelije je moguće formatirati i nekim unapred napravljenim stilovima
 - **Format as Table** – formatira skup ćelija kao da se radi o jednoj tabeli
 - **Cell Styles** – predefinisani stilovi koji su primenjivi na pojedinačne ćelije
 - **Conditional Formatting** – formatiranje ćelije na osnovu nekih uslova koji se definišu nad podacima



43

43

Formatiranje ćelije

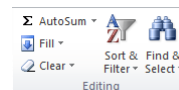
- **Conditional Formatting**
 - **Highlight Cell Rules**
 - zavisi od sadržaja ćelije
 - **Top/Bottom Rules**
 - zavisi od poređenja sadržaja ćelije sa sadržajem ostalih ćelija u opsegu u skladu sa nekim pravilom
 - **Data Bars, Color Scales, Icon Sets**
 - mogućnost ubacivanja stubića, ikonica i postepene promene boja pozadine ćelije u zavisnosti od sadržaja ćelija u opsegu

44

44

Uređivanje i filtriranje unetih podataka

- *Edit*-ovanje omogućava:
 - popunjavanje sadržaja ćelije
 - primenu nekih osnovnih funkcija na opseg
 - sumiranje, prosek, brojanje
 - popunjavanje ćelija u raznim pravcima praćenjem određenog šablona
 - sortiranje i filtriranje
 - pronalaženje određenog sadržaja



45

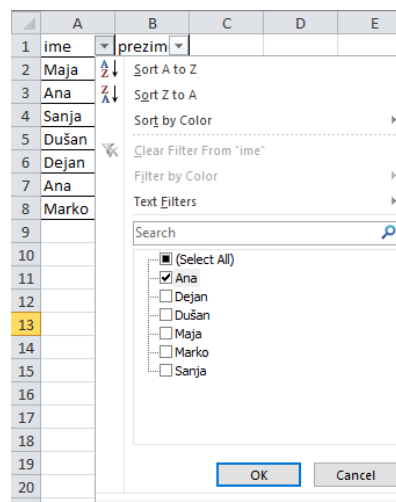
45

Uređivanje i filtriranje unetih podataka

- **Filteri** pružaju mogućnost jednostavnog pregleda, sortiranja i sužavanja velike tabele na osnovu jednog ili više zadatih kriterijuma
 - potrebno je označiti tabelu i uključiti filter iz opcije *Sort&Filter*

	A	B
1	ime	prezime
2	Maja	Perić
3	Ana	Milić
4	Sanja	Savić
5	Dušan	Vasić
6	Dejan	Marić
7	Ana	Topić
8	Marko	Ilić

	A	B
1	ime	prezime
3	Ana	Milić
7	Ana	Topić



46

46

Jednostavne matematičke funkcije

- Veći deo funkcija Excela služi za obradu grupa podataka
 - Funkcije rade nad opsegom (eng. *range*) ćelija
 - SUM(opseg1, opseg2, ...)
 ■ zbir vrednosti iz zadatih opsega
 - AVERAGE(opseg1, opseg2, ...)
 ■ aritmetička sredina vrednosti iz zadatih opsega
 - COUNT(opseg1, opseg2, ...)
 ■ broj numeričkih podataka, ili podataka koji se mogu svesti na numeričke, u zadatim opsezima
 - PRODUCT(opseg1, opseg2, ...)
 ■ proizvod vrednosti iz zadatih opsega
 - MIN(opseg1, opseg2, ...) i MAX(opseg1, opseg2, ...)
 ■ najmanja i najveća vrednost iz zadatih opsega

47

47

Jednostavne matematičke funkcije

Proizvodi	Cena po kg	Količina u kg	Cena		
višnje	200	2	400	=B2*C2	
trešnje	300	2.5	750		
jagode	150	1	150		
maline	500	0.5	250		
breskve	100	2	200		
kruške	130	1	130		
jabuke	80	3	240		
		Ukupno:	2120	=SUM(D2:D8)	
Broj proizvoda:	7	=COUNT(B2:B8)			
Najmanja cena:	80	=MIN(B2:B8)			
Najveća cena:	500	=MAX(B2:B8)			

48

48

Logičke funkcije

- Funkcije koje služe za proveru uslova
 - Logičke konstante TRUE i FALSE
 - AND(izraz1, izraz2, ...)
 - konjunkcija, TRUE, ako su vrednosti svih izraza tačne
 - OR(izraz1, izraz2, ...)
 - disjunkcija, FALSE, ako su vrednosti svih izraza netačne
 - NOT(izraz)
 - negacija
 - IF(uslov, vrednost_t, vrednost_n)
 - daje vrednost_t, ako je vrednost uslova TRUE, a vrednost_n, ako je vrednost uslova FALSE

49

49

Logičke funkcije

Izbori 2020.		Ukupan broj glasova	38776			
Partija	Broj glasova					
A	1345					
B	2345			Izbrojati glasove		
C	1237			=IF(F1<38000, "Ponoviti glasanje", "Izbrojati glasove")		
D	4321					
E	9654					
F	2345					
G	9875					
H	7654					

50

50

Funkcije

- Kombinovane funkcije
 - COUNTIF(opseg; kriterijum)
 - broj ćelija u opsegu koje zadovoljavaju zadati kriterijum
 - SUMIF(opseg; kriterijum; opseg_za_sumiranje)
 - sabiranje vrednosti u određenom opsegu koje zadovoljavaju zadati kriterijum
 - AVERAGEIF(opseg; kriterijum; opseg_za_računanje)
 - MAXIF, MINIF
 - ne postoje, moraju se kombinovati sa IF (nakon unosa formule treba pritisnuti *Ctrl+Shift+Enter*)
 - MAX(IF(logički_test; opseg))
 - MIN(IF(logički_test; opseg))

51

51

Funkcije

uspeh		fx =IF(G2>=4,5;"odličan";IF(G2>=3,5;"vrlo dobar";IF(G2>=2,5;"dobar";IF(G2>=1,5;"dovoljan";"nedovoljan"))))									
	A	B	C	D	E	F	G	H	I	J	K
1	Učenik	Matematika	Srpski	Engleski	Fizika	Informatika	Prosek				
2	Ana Jović	5	4	5	3	5	4,4	vrlo dobar			
3	Mila Savić	5	5	5	5	5	5	odličan			
4	Luka Marić	5	5	4	5	4	4,6	odličan			
5	Dejan Perić	3	3	4	4	3	3,4	dobar			
6	Marko Mitić	5	4	5	4	4	4,4	vrlo dobar			
7											
8								Broj odličih:		2	
9								Broj vrlo dobrih:		2	
10								Broj dobrih:		1	
11								Broj dovoljnih:		0	
12											
13		=AVERAGE(B2:F2)									
14		=IF(G2>=4,5;"odličan";IF(G2>=3,5;"vrlo dobar";IF(G2>=2,5;"dobar";IF(G2>=1,5;"dovoljan";"nedovoljan"))))									
15		=COUNTIF(H\$2:H\$6;"odličan")									
16											
17		Najveći prosek učenika koji su odlični:									
18		5 =MAX(IF(uspeh="odličan";prosek))									
19											

52

52

Funkcije

- Najveći prosek učenika koji su odlični
 - obeležiti opseg „uspeh“
 - =MAX(IF(uspeh="odličan";prosek))
 - *Ctrl+Shift+Enter*
 - poruka se odnosi na opseg i biće uokvirena vitičastim zagradama

53

53

Greške

- Ukoliko se prilikom unosa formule pojavi rezultat koji počinje znakom #, to znači da je došlo do greške prilikom izračunavanja
 - #####
 - širina ćelije nije dovoljna da se prikaže rezultat (potrebno je samo proširiti ćeliju)
 - #DIV/0!
 - pokušaj deljenja nulom ili praznom ćelijom
 - #NAME?
 - ne prepoznaje se ime funkcije koje se koristi u formuli
 - #N/A
 - podatak ćelije nije na raspolaganju

54

54

Greške

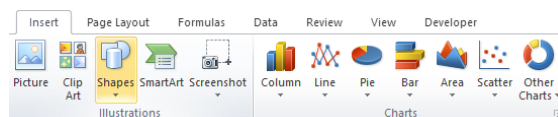
- Ukoliko se prilikom unosa formule pojavi rezultat koji počinje znakom #, to znači da je došlo do greške prilikom izračunavanja
 - **#NULL!**
 - formula koristi presek dva područja koja se ne seku (preseki područja odnose se na ćelije koje su zajedničke za oba područja)
 - **#NUM!**
 - postoji problem sa vrednošću (na primer, unet je negativan broj gde se očekuje pozitivan)
 - **#REF!**
 - u formulu je uneta najspravna referenca, koja se poziva za neispravnu adresu ćelije (na primer, ako je ćelija izbrisana sa radnog lista)
 - **#VALUE!**
 - formula sadrži operand pogrešnog tipa

55

55

Crtanje grafikona

- Podaci u radnom listu mogu se predstaviti grafički pomoću grafikona
- **Vrste grafikona**
 - **Stubičasti i trakasti grafikoni**
 - **Linijski grafikoni**
 - **Oblasni grafikoni**
 - **Kružni (pitasti) grafikoni**
 - **Prstenasti grafikoni**
 - **Radijalni grafikoni**
 - **XY grafikoni**
 - **Trodimenzionalni grafikoni**
 - **Površinski grafikoni**

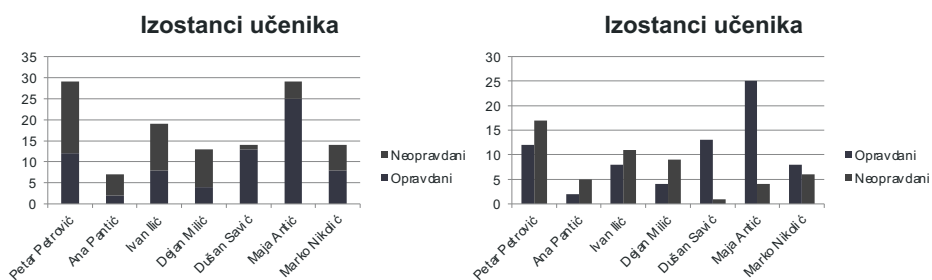


56

56

Stubičasti (eng. *column*) i trakasti (eng. *bar*) grafikoni

- Koriste se za poređenje vrednosti u različitim vremenskim periodima ili za poređenje stavki
- Trakasti grafikon se može posmatrati kao stubičasti grafikon položen na bok

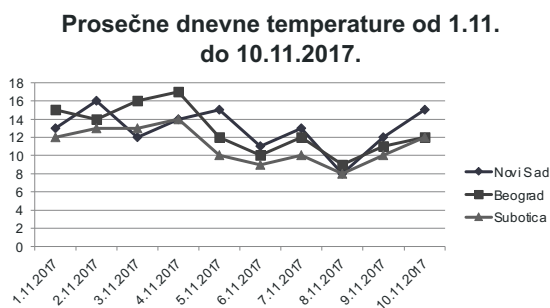


57

57

Linijski grafikoni (eng. *line*)

- Koristi se da ilustruje trend u ravnomernim vremenskim intervalima
 - Na primer, praćenje prosečne temperature po danima

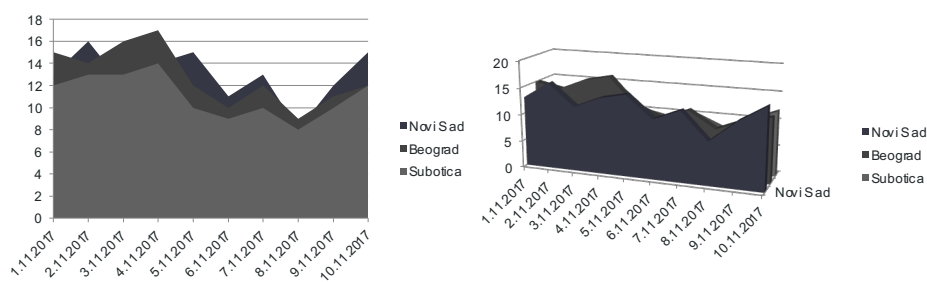


58

58

Oblasni grafikoni (eng. *area*)

- Linijski grafikoni u kojima je prostor ispod linije ispunjen



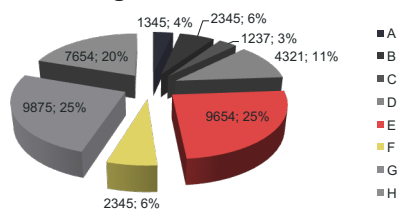
59

59

Kružni grafikoni (eng. *pie*)

- Obično se crta na osnovu jednog niza podataka
 - u jednoj koloni ili redu
- Pogodan za prikaz međusobnog odnosa pojedinih vrednosti prema celini

Procentualno učešće partija u glasovima



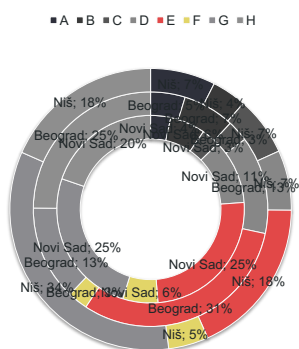
60

60

Prstenasti grafikon (eng. *doughnut*)

- Specijalna forma kružnog grafikona
 - odnos delova i celine, ali može sadržati više različitih grupa podataka

Procentualno učešće partija u glasovima



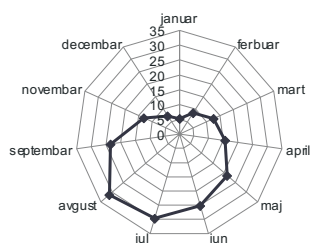
61

61

Radijalni grafikon (eng. *radar*)

- Grafikon i čijem je centru y osa
- Svaki tip podatka ima svoj krak (osu) koja izlazi iz centra
- Porede se ukupne vrednosti nekoliko različitih grupa podataka
 - npr. sezonske promene u nekom vremenskom intervalu

Maksimalne mesečne temperature u 2016.



62

62

XY grafikon (eng. *scatter*)

- Pogodan za prikaz odnosa između dva skupa vrednosti
 - pogotovo u nejednakim vremenskim intervalima
- Svaki pojedinačni podatak je koordinata koja se sastoji od jedne x i jedne y vrednosti

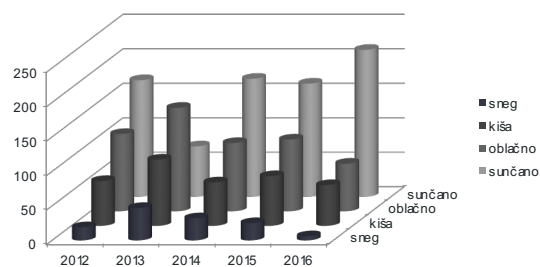


63

63

Trodimenzionalni grafikon

- Prikazuju se podaci u tri dimenzije
 - trakasti, stubičasti, kružni, piramidalni, ...

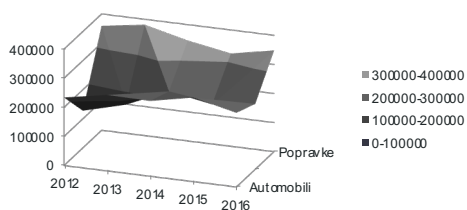


64

64

Površinski grafikon (eng. *surface*)

- Prikaz podataka koji zavise od dve veličine, na x i y osama

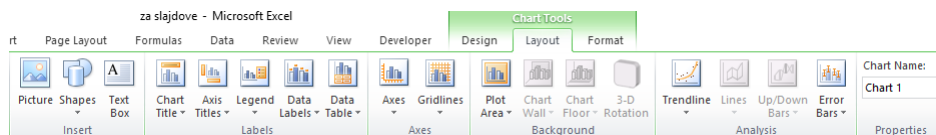


65

65

Grafikoni

- Podaci u radnom listu mogu se predstaviti grafički pomoću grafikona
- Podešavanje izgleda grafikona podrazumeva:
 - podešavanje naslova grafikona (*Chart Title*)
 - podešavanje naziva osa (*Axis Title*)
 - dodelu oznaka podacima (*Data Labels*)
 - prikazivanje ili skrivanje legende sa imenima serija podataka koji su prikazani na grafikonu (*Legend*)
 - prikazivanje ili skrivanje pomoćnih osa (*Gridlines*)



66

66

Grafikoni

- Podešavanja možemo raditi i na grafikonu

Učenik	Matematika	Srpski	Engleski	Fizika	Informatika	Prosek	
Ana Jović	5	4	5	3	5	4,4	vrlo dobar
Mila Savić	5	5	5	5	5	5	odličan
Luka Marić	5	5	4	5	4	4,6	odličan
Dejan Perić	3	3	4	4	3	3,4	dobar
Marko Mitić	5	4	5	4	4	4,4	vrlo dobar

Broj odličnih: 2
 Broj vrlo dobrih: 2
 Broj dobrih: 1
 Broj dovoljnih: 0

=AVERAGE(B2:F2)
 =IF(G2>=4,5,"odličan";IF(G2>=3,5,"vrlo dobar";IF(G2>=2,5,"dobar";IF(G2>=1,5,"dovoljan";"nedovoljan"))))
 =COUNTIF(H\$2:H\$6;"odličan")
 Najveći prosek učenika koji su odlični:
 =MAX(IF(uspeh="odličan";prosek))

67

Funkcije

- Funkcije
 - INT(vrednost)
 - vrši zaokruživanje naniže, ka najbližem celom broju
 - ROUND(vrednost; broj_decimala)
 - vrši zaokruživanje na željeni broj decimala
 - ROUNDUP(vrednost; broj_decimala)
 - zaokruživanje na prvi veći broj za zadati broj decimala
 - ROUNDDOWN(vrednost; broj_decimala)
 - zaokruživanje na prvi manji broj za zadati broj decimala
 - CEILING(vrednost; faktor)
 - zaokruživanje ka višem broju prema faktoru zaokruživanja
 - FLOOR(vrednost; faktor)
 - zaokruživanje ka nižem broju prema faktoru zaokruživanja

68

68

Funkcije

	A	B	C	D	E	F	G
1		=INT(A2)	=ROUND(A2;2)	=ROUNDUP(A2;4)	=ROUNDDOWN(A2;3)	=CEILING(A2;1)	=FLOOR(A2;1)
2	2,42356	2	2,42	2,4236	2,423	3	2
3	5,75312	5	5,75	5,7532	5,753	6	5
4	1,21711	1	1,22	1,2172	1,217	2	1
5	3,81171	3	3,81	3,8118	3,811	4	3

69

69

Funkcije

- Funkcije
 - POWER(broj; stepen)
 - stepen broja
 - SQRT(broj)
 - kvadratni koren broja
 - ABS(broj)
 - apsolutna vrednost broja
 - PI()
 - vrednost broja Pi

16	=POWER(2;4)	
5	=SQRT(25)	
2	=ABS(-2)	
3,141593	=PI()	

70

70

Funkcije

- Funkcije
 - LEFT(tekst; broj)
 - izdvajanje prvih (broj) nekoliko znakova iz teksta
 - RIGHT(tekst; broj)
 - izdvajanje poslednjih (broj) nekoliko znakova iz teksta
 - MID(tekst; start; broj)
 - izdvajanje određenog broja znakova iz teksta počevši od zadate pozicije
 - FIND(string; tekst; start) i SEARCH(string; tekst; start)
 - pronalaženje pozicije od koje počinje neki string unutar zadatog teksta
 - FIND pravi razliku između malih i velikih slova

71

71

Funkcije

- Funkcije
 - LOWER(tekst), UPPER(tekst)
 - pretvaranje velikih slova u mala i obrnuto
 - CONCATENATE(tekst1, tekst2, ...)
 - spajanje tekstova; može da se koristi i znak &
 - LEN(tekst)
 - dužina teksta
 - SUBSTITUTE(tekst, stari_tekst, novi_tekst, broj_pojave)
 - zamena starog_teksta novim_tekstom u tekstu
 - PROPER(tekst)
 - pretvaranje početnih slova u rečima u velika
 - TRIM(tekst)
 - uklanjanje viška razmaka iz teksta

72

72

Funkcije za rad sa tekstom

Ovo je neki tekst.				
Ovo	=LEFT(A1;3)			
st.	=RIGHT(A1;3)			
o je	=MID(A1;3;4)			
	5 =FIND("je";A1;1)			
ovo je neki tekst.	=LOWER(A1)			
OVO JE NEKI TEKST.	=UPPER(A1)			
	18 =LEN(A1)			
Ovo je neki tekst.Ovo je neki tekst.	=CONCATENATE(A1;A1)			
Ovo jx nxki txkst.	=SUBSTITUTE(A1;"e";"x")			
Ovo Je Neki Tekst.	=PROPER(A1)			
Ovo je neki tekst.				
Ovo je neki tekst.	=TRIM(A18)			

73/36

73

Funkcije

- Informacione funkcije
 - ispituju da li ćelije zadovoljavaju neki uslov?
 - ispituju tip podatka u ćeliji
 - ISEVEN(broj)
 - TRUE, ako je broj paran, FALSE, ako nije
 - ISODD(broj)
 - TRUE, ako je broj neparan, FALSE, ako nije
 - ISNUMBER(vrednost)
 - da li je vrednost broj
 - ISERROR
 - da li je u izračunavanju nastupila greška
 - COUNTBLANK
 - koliko ima praznih ćelija u opsegu

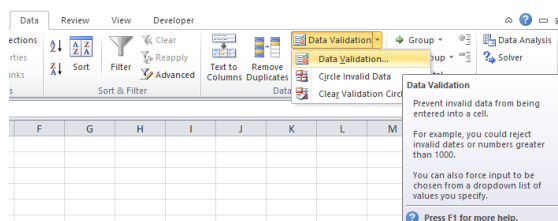
2	TRUE	=ISEVEN(G7)
a	FALSE	=ISODD(G7)
	TRUE	=ISNUMBER(G7)
	FALSE	=ISNUMBER(G8)

74

74

Provera valjanosti podataka

- *Data* → *Data Validation*
 - postavljanje ograničenja u vezi sa podacima koji se unose u ćeliju
 - doprinosi smanjivanju grešaka prilikom ručnog unosa
 - moguće je definisati upozorenja u slučaju unosa pogrešnih podataka, kao i uputstva o ograničenjima

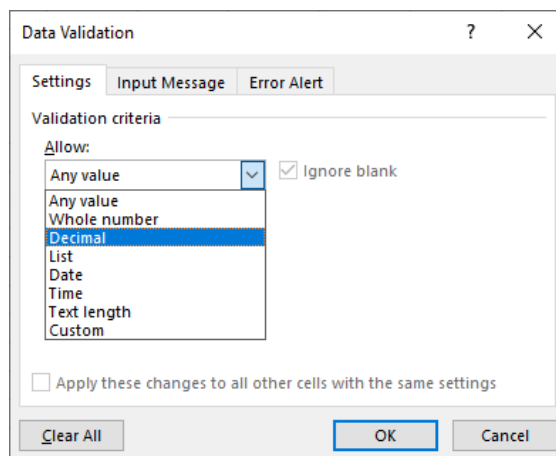


75

75

Provera valjanosti podataka

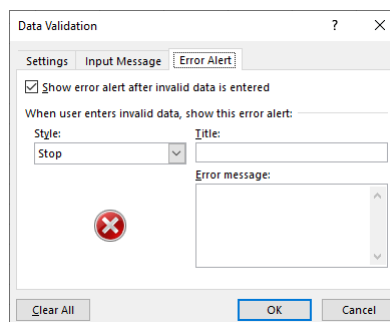
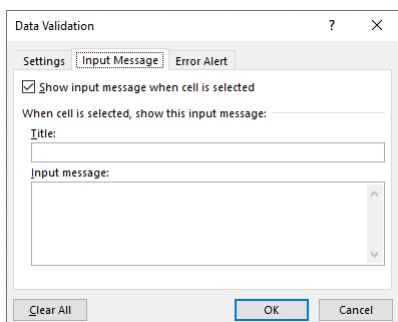
- Tipovi ograničenja:
 - ograničenje podataka u unapred definisanim stavkama na listi
 - ograničenje brojeva u određenom ospegu
 - ograničenje datuma u određenom vremenskom okviru
 - ograničenje vremena izvan određenog vremenskog okvira
 - ograničenje broja tekstualnih znakova
 - itd.



76

76

Provera valjanosti podataka

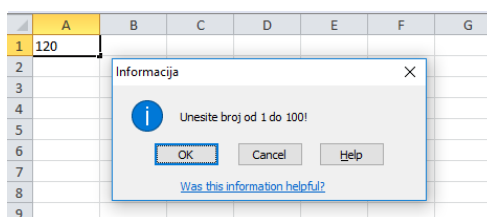
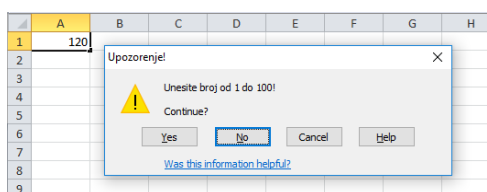
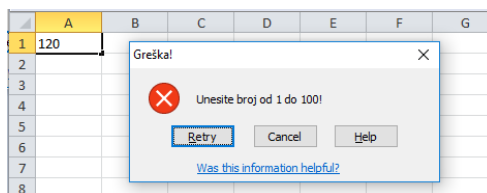


77

77

Provera valjanosti podataka

- Postoje tri tipa upozorenja:
 - **Stop**
 - sprečava korisnika da unese pogrešne podatke u ćeliju
 - **Warning**
 - upozorava korisnika da su uneti podaci pogrešni, bez sprečavanja unosa tih podataka
 - **Information**
 - obaveštava korisnika da su podaci koji su uneti nevalidni bez sprečavanja unosa tih podataka (ovaj tip upozorenja o grešci je najfleksibilniji)



78

78

Statističke funkcije

- **Statistika je oblast matematike koja se bavi sakupljanjem, analizom, interpretacijom, objašnjavanjem i prezentacijom velike količine podataka**
- Statistika izučava masovne pojave u društvu, tehnici i prirodi
 - kod masovnih pojava karakteristično je da pojedini slučajevi više ili manje odstupaju od onoga što se smatra karakteristikom
- **Populacija** – skup svih elemenata, osnovni skup ili generalna kolekcija
 - broj ljudi u zemlji ili gradu ili regiji
 - broj automobila proizvedenih u nekoj godini, svi studenti na univerzitetu
 - po pravilu ima veliki broj elemenata i može biti i beskonačna

79

79

Statističke funkcije

- **Osnovni zadatak statistike je definisanje raspodele verovatnoće**
- Uzorak je konačan podskup populacije
 - ne može se očekivati potpuno tačno opisivanje ili reprezentacija populacije
 - mora biti reprezentativan
 - svaki element populacije treba da ima podjednaku šansu da se nađe u uzorku
- **Zadatak statističke analize je:**
 - prikupljanje podataka
 - prikazivanje podataka
 - analiza podataka
 - zaključivanje na osnovu podataka

80

80

Statističke funkcije

- Statističke funkcije
 - Služe za izvođenje statističkih analiza nad nizovima podataka
 - AVERAGE(broj1; broj2; ...)
 - aritmetička sredina argumenata
 - MODE(broj1; broj2; ...)
 - broj koji se najčešće pojavljuje, odnosno ponavlja u zadatom opsegu
 - MEDIAN(broj1; broj2; ...)
 - broj koji se nalazi u sredini zadanog opsega brojeva; polovina brojeva ima manju vrednost od njega, a polovina veću
 - STDEVP(broj1; broj2; ...)
 - standardna devijacija, u statistici apsolutna mera disperzije u osnovnom skupu
 - koliko u proseku elementi skupa odstupaju od aritmetičke sredine skupa

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}$$

gde je:

N - broj elemenata u skupu

μ - aritmetička sredina skupa

x_i - i -ti član skupa ($i = 1, 2, \dots, N$)

81

81

Statističke funkcije - AVERAGE

		Predmeti					Prosečna ocena učenika
		Matematika	Fizika	Hemija	Biologija	Geografija	
Učenici	Petar Petrović	2	4	4	4	4	3.6
	Ana Pantić	3	3	2	2	3	2.6
	Ivan Ilić	4	2	5	1	4	nedovoljan
	Dejan Milić	1	3	4	2	3	nedovoljan
	Dušan Savić	2	4	3	4	2	3
	Maja Antić	3	1	2	3	3	nedovoljan
	Marko Nikolić	5	3	3	5	5	4.2
		=IF(COUNTIF(C41:G41,1)>0,"nedovoljan",AVERAGE(C41:G41))					

82

82

Statističke funkcije – MEDIAN i MODE

Studenti	Ocene na ispitu	
Petar	9	
Milan	8	
Sava	9	
Ana	7	
Marko	10	
Sanja	8	
Iva	9	
Dunja	9	
Filip	10	
Darko	7	
Maja	10	
Najčešće ponavljana ocena	9	=MODE(W26:W36)
Srednja ocena	8.727272727	=AVERAGE(W26:W36)
Median	9	=MEDIAN(W26:W36)

83

83

Statističke funkcije - MEDIAN

Studenti	Ocene na ispitu		
Ana	7		
Darko	7		
Milan	8		
Sanja	8		
Petar	9		
Sava	9		
Iva	9		
Dunja	9		
Marko	10		
Filip	10		
Maja	10		
		Ukupno 11 brojeva	
		Na srednjoj (šestoj)	
		poziciji nalazi se	
		broj 9	

84

84

Statističke funkcije - MEDIAN

	Plate		Plate
Petar	45000	Iva	35500
Milan	44000	Sanja	38000
Sava	51000	Dunja	39000
Ana	56000	Filip	41000
Marko	43000	Marko	43000
Sanja	38000	Milan	44000
Iva	35500	Petar	45000
Dunja	39000	Sava	51000
Filip	41000	Ana	56000
Darko	125000	Darko	125000
Maja	250000	Maja	250000
Srednja	69772.73	=AVERAGE(J49:J59)	
Median	44000	=MEDIAN(J49:J59)	

85

85

Statističke funkcije - STDEVP

Studenti	Ocene na ispitu				
Petar	9	0.272727273	=W70-\$W\$82	0.07438	=X70*X70
Milan	8	-0.727272727		0.528926	
Sava	9	0.272727273		0.07438	
Ana	7	-1.727272727		2.983471	
Marko	10	1.272727273		1.619835	
Sanja	8	-0.727272727		0.528926	
Iva	9	0.272727273		0.07438	
Dunja	9	0.272727273		0.07438	
Filip	10	1.272727273		1.619835	
Darko	7	-1.727272727		2.983471	
Maja	10	1.272727273		1.619835	
Najčešće ponavljana ocena	9	=MODE(W26:W36)			
Srednja ocena	8.727272727	=AVERAGE(W26:W36)			
Median	9	=MEDIAN(W26:W36)			
St. devijacija	1.052348809			1.052349	=SQRT(SUM(Z70:Z80)/COUNT(Z70:Z80))

86

86

Statističke funkcije

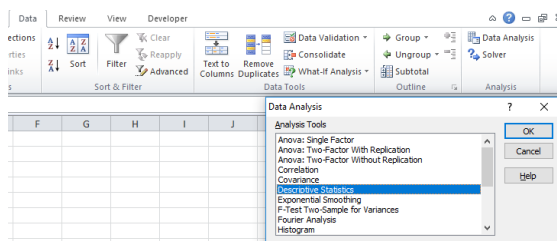
- Statističke funkcije
 - VARIANCE(broj1; broj2; ...)
 - računanje varijanse koja predstavlja standardnu devijaciju na kvadrat
 - CORREL(niz1; niz2; ...)
 - računanje koeficijenta korelacije koji služi za poređenje dva skupa brojeva
 - kako bi se utvrdilo koliko su oni slični
 - rezultat je vrednost između -1 i 1; što je vrednost bliža 1, to su i skupovi sličniji
 - VARP(broj1; broj2; ...)
 - kovarijansa određuje meru zajedničkog varijabiliteta dve promenljive

87

87

Deskriptivna statistika

- *Data* → *Analysis* → *Data Analysis* → *Descriptive Statistics*
 - korišćenjem opcije Descriptive Statistics dobija se preko 10 različitih statističkih funkcija
 - loša strana korišćenja ove opcije je što su dobijene vrednosti **statičke**
 - ne menjaju se promenom vrednosti u izvornim podacima ne menjaju se vrednosti izračunatih statistika



88

88

Deskriptivna statistika

- **Standard Error**
 - standardna devijacija svih uzoračkih sredina
- **Kurtosis**
 - mera homogenosti raspodele
- **Skewness**
 - mera asimetrije raspodele
 - ukoliko je vrednost 0, raspodela je simetrična
- **Range**
 - razlika između najveće i najmanje vrednosti

	A	B	C	D	E
1	Plate za oktobar				Column1
2	23456				
3	24564			Mean	62069,66667
4	45324			Standard Error	13212,41086
5	51432			Median	45324
6	23654			Mode	#N/A
7	65434			Standard Deviation	51171,44721
8	76533			Sample Variance	2618517010
9	43234			Kurtosis	4,866857786
10	21345			Skewness	2,042190023
11	76545			Range	191111
12	34523			Minimum	21345
13	21435			Maximum	212456
14	87654			Sum	931045
15	123456			Count	15
16	212456				
17					

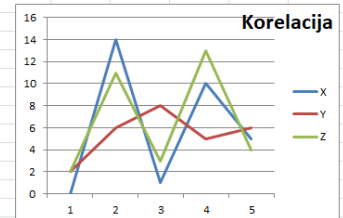
89

89

Korelacija

- Koeficijent korelacije pruža uvid koliko su jako povezane dve promenljive
 - **ima vrednost između -1 i 1**
 - vrednost 1 implicira postojanje savršene pozitivne korelacije (veze) – ukoliko vrednost promenljive X raste, raste i vrednost promenljive Y i obrnuto
 - vrednost -1 implicira postojanje savršene negativne korelacije (veze) – ukoliko vrednost promenljive X raste, vrednost promenljive Y opada i obrnuto
 - vrednost 0 implicira da ne postoji korelacija

	A	B	C	D	E	F	G	H	I
1	X	Y	Z				X	Y	Z
2	0	2	2			X	1		
3	14	6	11			Y	0,191516	1	
4	1	8	3			Z	0,909268	0,108893	1
5	10	5	13						
6	5	6	4						
7									
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									



90

90

Kovarijansa

- Kovarijansa meri kako promene jedne promenljive utiču na drugu
 - pozitivna vrednost kovarijanse implicira da su promenljive pozitivno povezane
 - negativna vrednost kovarijanse implicira da su promenljive negativno povezane
 - velika vrednost kovarijanse implicira da postoji jaka veza između promenljivih, dok mala vrednost implicira slabu vezu

91

91

Pivot tabele

- **Pivot tabele predstavljaju alat za kreiranje interaktivnih izveštaja koji se prave na osnovu selektovanih podataka ili tabela**
 - kolone iz izvornih tabela se koriste kao dimenzije izveštavanja (vrši se sabiranje brojeva u njima) ili se koriste za analizu, kao filteri
 - napravljene Pivot tabele se lako menjaju dodavanjem ili uklanjanjem dimenzija
 - više različitih pogleda na jedinstven skup podataka

92

92

Pivot tabele

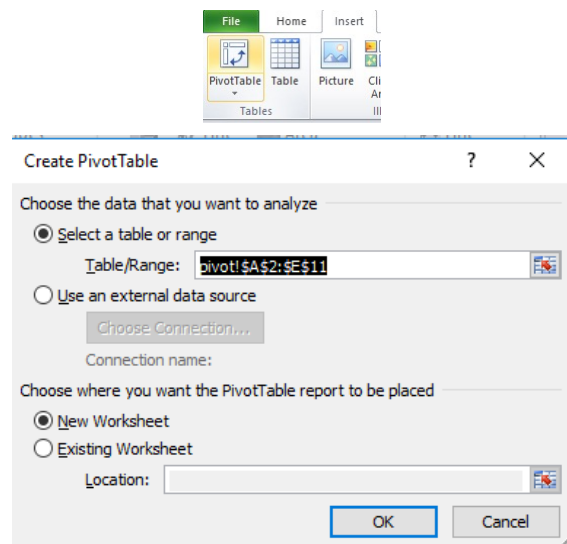
- Za kreiranje Pivot tabele potrebno je imati dobro pripremljene podatke
 - moraju se nalaziti unutar strukturirane tabele ili standardne Excel tabele
 - tabela mora imati zaglavlje gde su opisani podaci
 - na primer, IME, PREZIME, PLATA, ...
- Korišćenjem Pivot tabele mogu se dobiti unakrsni i vrlo detaljni međuzbirovi
 - dobijene informacije mogu se raspoređivati na razne načine

93

93

Pivot tabele

- Kreiranje pivot tabele počinje selektovanjem željenih podataka (opseg podataka, strukturirana tabela, imenovani blok ćelija)
 - u polje *Pivot/Range* potrebno je uneti opseg ćelija početne tabele
 - ukoliko se koriste podaci iz spoljašnje datoteke (na primer, baze podataka) bira se opcija *Use an external data source*
 - potrebno je odrediti da li se *Pivot* tabela kreira u novom radnom listu ili na određenoj poziciji u postojećem radnom listu
 - opcija *Add Data to Data Model* (od verzije Excel 2013) odnosi se na mogućnost dodavanja tabele u model podataka

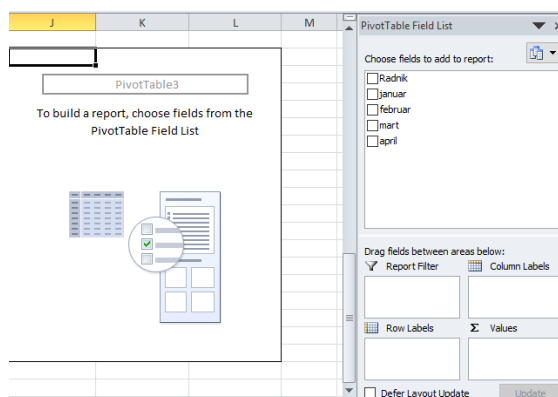


94

94

Pivot tabelle

- Nakon potvrde kreiranja, na odabranoj poziciji u radnom listu pojaviti se objekt Pivot table, kao i kontrolni panel sa desne strane lista, *Pivot Table Fields*
 - u gornjem delu panela nalaze se polja, odnosno dimenzije izveštavanja, a u donjem delu nalaze se četiri oblasti: *Values*, *Rows*, *Columns* i *Filters*
 - selektovanjem ili prevlačenjem dimenzija u neku od ove četiri oblasti kreiraju se izveštaji
 - u slučaju selektovanja Excel sam zaključuje u kojoj oblasti bi neka dimenzija trebalo da se pojavi, dok prevlačenjem korisnik sam bira gde će se ona nalaziti

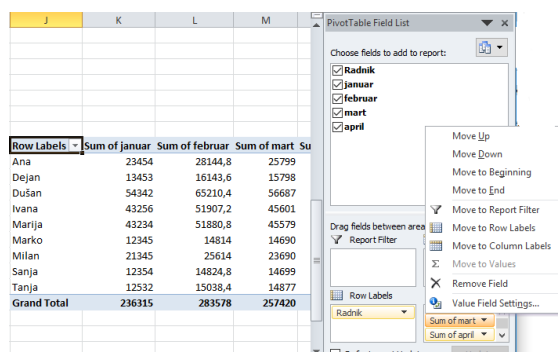


95

95

Pivot tabelle

- U oblasti vrednosti (Values) treba da se nalaze one kolone nad kojima se vrše neka izračunavanja
- U redovima (Rows) ili kolonama (Columns) nalaze se one dimenzije po kojima se vrši izračunavanje (sumiranje, prebrojavanje, ...)
- Filteri (Filters) predstavljaju one dimenzije po kojima se filtriraju podaci u tabeli
- U polju Values za izabrane podatke moguće je promeniti funkciju
 - selektovati podatak kojem se želi promeniti funkcija, pa izabrati opciju Value Field Settings
 - nakon toga u prozoru koji se pojavljuje može se izabrati i promeniti funkcija



96

96

Pivot table

	A	B	C	D	E	F
1	Radnik	Mesec	Broj sati		Row Labels	Sum of Broj sati
2	Marko	januar	23		Ana	69
3	Milan	januar	13		Dejan	37
4	Ana	januar	32		Dušan	34
5	Dušan	januar	14		Ivana	83
6	Sanja	januar	35		Marija	37
7	Ivana	januar	25		Marko	70
8	Dejan	januar	12		Milan	44
9	Tanja	januar	24		Sanja	52
10	Marija	januar	16		Tanja	34
11	Marko	februar	17		Grand Total	460
12	Ana	februar	25			
13	Ivana	februar	18			
14	Dejan	februar	25			
15	Milan	februar	31			
16	Dušan	mart	20			
17	Ana	mart	12			
18	Sanja	mart	17			
19	Ivana	mart	40			
20	Marko	mart	30			
21	Marija	mart	21			
22	Tanja	mart	10			
23						

PivotTable Field List

Choose fields to add to report:

Radnik
 Mesec
 Broj sati

Drag fields between areas below:

Report Filter Column Labels

Row Labels: Radnik
 Values: Sum of Broj sati

Defer Layout Update Update

97/36

97

Pivot table

- Moguće je podatke iz pivot table predstaviti i na grafikonu

File Home Insert
PivotTable Table Picture CI
PivotTable PivotChart

Total

Sort (Radnik)

Sort options

Manual (you can drag items to rearrange them)

Ascending (A to Z) by: Sum of Broj sati

Descending (Z to A) by: Sum of Broj sati

Summary
Sort Radnik by Sum of Broj sati in ascending order

Choose fields to add to report:

Radnik
 Mesec
 Broj sati

Drag fields between areas below:

Report Filter Legend Fields ...

Axis Fields (Ca... Values

Radnik Sum of Broj sati

Defer Layout Update Update

98

98

Funkcije za pretraživanje i reference (*Lookup and reference functions*)

- Funkcije za pretraživanje i reference
 - Koriste se kada postoje tabele koje su u međusobnoj vezi, tj. postoji tabela u kojoj se nalaze određeni podaci koje treba iskoristiti u nekoj drugoj tabeli
 - **CHOOSE**(indeks; v1; v2; ...)
 ■ vraća jednu od najviše 29 vrednosti v_i , u zavisnosti od vrednosti indeksa
 - **LOOKUP**(lookup_vrednost; lookup_vektor)
 ■ pretražuje jedan red ili kolonu i pronalazi vrednost sa iste pozicije u drugom redu ili koloni
 - **VLOOKUP**(lookup_vrednost; tabela; indeks_kolone; [lookup_opseg])
 ■ pronalazi red u tabeli koji u prvoj koloni sadrži traženu vrednost
 ■ rezultat je sadržaj ćelije iz kolone koja je zadata indeksom
 ■ podaci treba da budu organizovani tako da vrednost koja se traži bude sa leve strane vrednosti koja treba da bude vraćena
 ■ ako lookup_opseg ima vrednost FALSE, traži se potpuno podudaranje

99

99

Funkcije za pretraživanje i reference - VLOOKUP

ŠifraJela			
	A	B	C
1	Šifra	Naziv	Cena
2	1001	Čevapi	300.00
3	1002	Vešalica	225.00
4	1003	Ražnjići	250.00
5	1004	Sok 0.25	40.00
6	1005	Pivo 0.33	55.00
7	1006	Pivo 0.5	50.00
8	1007	Vino 0.5	100.00
9	1008	Vino 1l	300.00
10	1009	Kolači	60.00
11	1010	Palačinke	52.00
12	1011	Sladoled	40.00
13	1012	Hleb	15.00
14	1013	Salata	25.00

Restoran Lav					
Bulevar cara Lazara 12, Telefon; 344-324					
Br.	Šifra	Količina	Naziv	Jed. cena	Cena
1	1001	2	Čevapi	300.00	600.00
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
Ukupno:					600.00

Šifre jela
 Unesite šifru
 jela iz šifarnika
 ili je unesite

=IF(ISNUMBER(B9), VLOOKUP(B9, Jela, 2, FALSE), "")
 =IF(ISNUMBER(B9), VLOOKUP(B9, Jela, 3, FALSE), "")
 =IF(AND(ISNUMBER(C9), ISNUMBER(E9)), C9*E9, "")

100/36

100

Funkcije za rad sa bazama podataka (Database Functions)

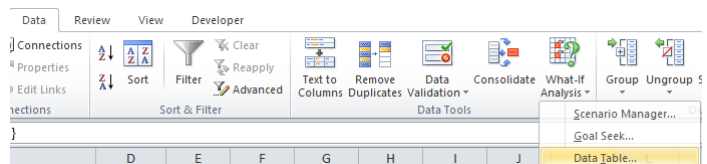
- Funkcije za rad sa bazama podataka
 - Koriste se za ispitivanje da li podaci u listi zadovoljavaju neke uslove
 - lista, odnosno baza podataka je tabela srodnih podataka
 - DCOUNT(baza, polje, kriterijum)
 - prebrojava ćelije koje sadrže brojeve koji zadovoljavaju dati kriterijum i nalaze se u zadatom polju baze podataka, odnosno koloni liste
 - DMAX, DMIN, DSUM
 - DGET(baza, polje, kriterijum)
 - izdvaja jednu vrednost iz zadate kolone liste, odnosno polja baze podataka, koja pripada slogu koji zadovoljava zadati kriterijum

101

101

What-if analiza

- **What-if (šta-ako) analiza je proces promene vrednosti u ćelijama kako bi se video uticaj tih promena na krajnji ishod**
 - analiza različitih scenarija – šta bi se desilo kada bi se neki parametar promenio
- Postoje tri vrste alata za „šta-ako“ analizu:
 - upravljanje scenarijom (*Scenario Manager*)
 - „traženje cilja“ (*Goal Seek*)
 - tabele sa podacima (*Data Table*)
- *Data* → *Data Tools*



102

102

What-if analiza

- Upravljanje scenarijom omogućava da se analiziraju sve moguće vrednosti jedne funkcije kada se promeni jedan parametar
 - koliko bi iznosio profit ukoliko bi bilo prodato 60% knjiga po višoj ceni, a ako bi bilo prodato 70% knjiga po višoj ceni?

103

103

What-if analiza

	A	B	C	D	E
1	Knjižara				
2		ukupan broj knjiga	% prodato po višoj ceni		
3		100	60%		
4					
5					
6			broj knjiga	cena po knjizi	
7		viša cena	60	\$50	
8		niža cena	40	\$20	
9					
10			ukupna zarada	\$3,800	
11					
12					
13					
14					

Scenario Summary						
Current Values:	60% po višoj ceni	70% po višoj ceni	80% po višoj ceni	90% po višoj ceni	100% po višoj ceni	
Changing Cells:						
\$C\$4	60%	60%	70%	80%	90%	100%
Result Cells:						
\$D\$10	3,800 RSD	3,800 RSD	4,100 RSD	4,400 RSD	4,700 RSD	5,000 RSD

Notes: Current Values column represents values of changing cells at time Scenario Summary Report was created. Changing cells for each scenario are highlighted in gray.

104/36

104

What-if analiza

- Traženje cilja (*Goal Seek*) pokazuje koliko treba da se promeni vrednost jedne promenljive kako bi se dobio unapred određeni nivo ciljne promenljive
 - koliko knjiga po višoj ceni treba prodati da bi bio ostvaren profit od 4400 dinara?
 - koliki treba da bude broj meseci otplate da bi rata kredita bila 9000 dinara?
 - kolika treba da bude kamatna stopa da bi rata kredita bila 9000 dinara?
- *Set Cell* – adresa ćelije koja se menja
- *To Value* – vrednost na koju ona treba da se promeni
- *By Changing cell* – adresa ćelije koja se menja, a koja utiče na konačan ishod

105

105

What-if analiza

D10		fx		=C7*D7+C8*D8			
A	B	C	D	E	F	G	H
1	Knjižara						
2							
3	ukupan broj knjiga	% prodat po višoj ceni					
4	100	60%					
5							
6			broj knjiga	cena po knjizi			
7	viša cena	60	\$50				
8	niža cena	40	\$20				
9							
10			ukupna zarada	\$3.800			
11							
12							
13							

Goal Seek dialog box:

Set cell: D10
To value: 4400
By changing cell: \$C\$4

D10		fx		=C7*D7+C8*D8			
A	B	C	D	E	F	G	H
1	Knjižara						
2							
3	ukupan broj knjiga	% prodat po višoj ceni					
4	100	80%					
5							
6			broj knjiga	cena po knjizi			
7	viša cena	80	\$50				
8	niža cena	20	\$20				
9							
10			ukupna zarada	\$4.400			
11							
12							
13							

Goal Seek Status dialog box:

Goal Seeking with Cell D10 found a solution.
Target value: 4400
Current value: \$4.400

106

What-if analiza

The screenshots illustrate the process of using Goal Seek in Excel to find a specific value for a formula by changing one of its input cells. In this case, the goal is to achieve a monthly payment (Rata) of 9,000.00 RSD by adjusting the interest rate (Kamatna stopa) from 5,00% to 7,02%.

107

107

What-if analiza

- Tabele podataka (*Data Table*) pružaju mogućnost da argumenti formule budu preuzeti iz redova ili kolona specijalno oformljene tabele
 - koristi se za funkcije koje imaju više argumenata, kako bi se videla vrednost rezultata u zavisnosti od promene argumenata
 - može se kreirati tabela sa jednom promenljivom i tabela sa dve promenljive

108

108

What-if analiza

Mesečna rata za kredit	
Pozajmica	2.500.000 RSD
Kamata	5%
Broj meseci otplate	480
Mesečna rata	12.054,92 RSD

Pozajmica	Mesečna rata
120	26516,37881
150	22447,62777
180	19769,84067
210	17886,38995
240	16498,89348
270	15441,52729
300	14614,75104
330	13955,18351
360	13420,54058
390	12981,52496
420	12617,19186
450	12312,16826
480	12054,91501

109/36

109

What-if analiza

Mesečna rata za kredit		Kamatna stopa									
Mesečna rata		4,50%	4,75%	5,00%	5,25%	5,50%	5,75%	6,00%	6,25%	6,50%	
12.054,92 RSD											
1.000.000 RSD	4495,628	4657,579	4821,966	4988,703	5157,703	5328,876	5502,136	5677,396	5854,568		
1.250.000 RSD	5619,535	5821,973	6027,458	6235,879	6447,129	6661,095	6877,671	7096,745	7318,21		
1.500.000 RSD	6743,443	6986,368	7232,949	7483,055	7736,554	7993,315	8253,205	8516,094	8781,852		
1.750.000 RSD	7867,35	8150,762	8438,441	8730,231	9025,98	9325,534	9628,739	9935,443	10245,49		
2.000.000 RSD	8991,257	9315,157	9643,932	9977,407	10315,41	10657,75	11004,27	11354,79	11709,14		
2.250.000 RSD	10115,16	10479,55	10849,42	11224,58	11604,83	11989,97	12379,81	12774,14	13172,78		
2.500.000 RSD	11239,07	11643,95	12054,92	12471,76	12894,26	13322,19	13755,34	14193,49	14636,42		
2.750.000 RSD	12362,98	12808,34	13260,41	13718,93	14183,68	14654,41	15130,88	15612,84	16100,06		
3.000.000 RSD	13486,89	13972,74	14465,9	14966,11	15473,11	15986,63	16506,41	17032,19	17563,7		
3.250.000 RSD	14610,79	15137,13	15671,39	16213,29	16762,53	17318,85	17881,94	18451,54	19027,35		
3.500.000 RSD	15734,7	16301,52	16876,88	17460,46	18051,96	18651,07	19257,48	19870,89	20490,99		

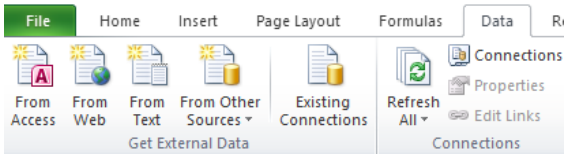
		Kamatna stopa									
Broj meseci otplate	Mesečna rata	4,50%	4,75%	5,00%	5,25%	5,50%	5,75%	6,00%	6,25%	6,50%	
150	21821,74	22133,38	22447,63	22764,47	23083,89	23405,88	23730,43	24057,52	24387,15		
180	19124,83	19445,8	19769,84	20096,94	20427,09	20760,25	21096,42	21435,57	21777,68		
210	17222,42	17552,64	17886,39	18223,64	18564,35	18908,51	19256,08	19607,04	19961,34		
240	15816,23	16155,59	16498,89	16846,1	17197,18	17552,09	17910,78	18273,21	18639,33		
270	14740,52	15088,86	15441,53	15798,48	16159,66	16525,02	16894,49	17268,01	17645,53		
300	13895,81	14252,93	14614,75	14981,19	15352,19	15727,66	16107,54	16491,73	16880,18		
330	13218,76	13584,46	13955,18	14330,83	14711,32	15096,54	15486,42	15880,84	16279,71		
360	12667,13	13041,18	13420,54	13805,09	14194,73	14589,32	14988,76	15392,93	15801,7		
390	12211,66	12593,81	12981,52	13374,66	13773,08	14176,64	14585,2	14998,62	15416,75		
420	11831,42	12221,42	12617,19	13018,58	13425,41	13837,52	14254,74	14676,91	15103,86		
450	11511,06	11908,64	12312,17	12721,46	13136,32	13556,55	13981,97	14412,39	14847,6		
480	11239,07	11643,95	12054,92	12471,76	12894,26	13322,19	13755,34	14193,49	14636,42		

110

110

Uvoz podataka

- Podaci se mogu uvesti iz različitih izvora
 - tekstualni fajlovi
 - Microsoft Access
 - SQL server
 - Oracle DB
 - ...



Uvoz podataka

Text Import Wizard - Step 1 of 3

The Text Wizard has determined that your data is Delimited. If this is correct, choose Next, or choose the data type that best describes your data.

Original data type

Choose the file type that best describes your data:

- Delimited - Characters such as commas or tabs separate each field.
- Fixed width - Fields are aligned in columns with spaces between each field.

Start import at row: 1 File origin: 852 - Central European (DOS)

Preview of file C:\Nastava\FTN - Uvod u informacioni i finansijski inzenjering\Excel\napre... \mena.txt

1 Ana Peric 34000
2 Milan Savic 35000
3 Dusan Maric 43000
4 Tanja Ilic 41000
5 Marko Vasic 43000

Text Import Wizard - Step 2 of 3

This screen lets you set the delimiters your data contains. You can see how your text is affected in the preview below.

Delimiters

- Tab
- Semicolon
- Comma
- Space
- Other

Treat consecutive delimiters as one

Text qualifier: *

Data preview

Ana Peric 34000
Milan Savic 35000
Dusan Maric 43000
Tanja Ilic 41000
Marko Vasic 43000

Text Import Wizard - Step 3 of 3

This screen lets you select each column and set the Data Format.

Column data format

- General
- Text
- Date: DMY
- Do not import column (skip)

General converts numeric values to numbers, date values to dates, and all remaining values to text.

Advanced Text Import Settings

Settings used to recognize numeric data

Decimal separator: .

Thousands separator: .

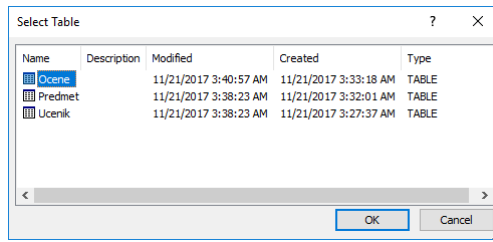
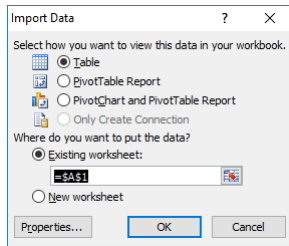
Note: Numbers will be displayed using the numeric settings specified in the Regional Settings control panel.

Trailing minus for negative numbers

Excel spreadsheet preview:

	A
1	Ana Peric 34000
2	Milan Savic 35000
3	Dusan Maric 43000
4	Tanja Ilic 41000
5	Marko Vasic 43000
6	Petar Rakic 51000
7	Maja Arsic 48000
8	Anja Milic 45000
9	Sava Popic 47000
10	Mila Dakic 54000
11	Ivana Ivic 46000
12	
13	

Uvoz podataka

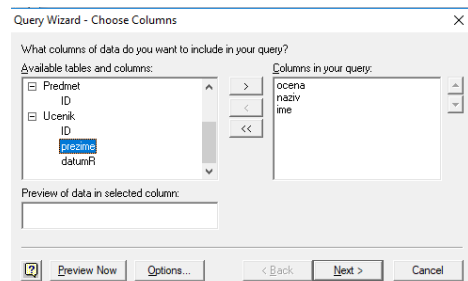
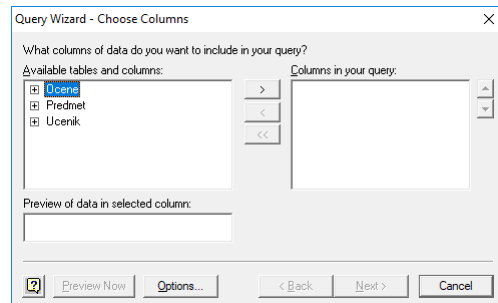
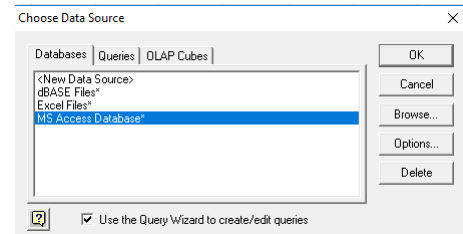
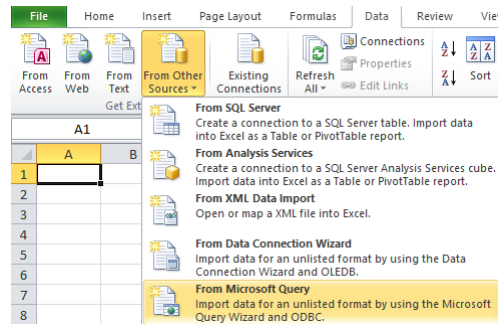


	A	B	C	D
1	ID	id_ucenika	id_predmeta	ocena
2	1	1	1	5
3	2	1	1	5
4	3	1	1	4
5	4	1	1	5
6	5	1	2	4
7	6	1	2	4
8	7	1	2	3
9	8	1	3	5
10	9	1	3	5
11	10	1	3	4
12	11	2	1	4
13	12	2	1	4
14	13	2	1	5
15	14	2	2	3
16	15	2	2	2

113

113

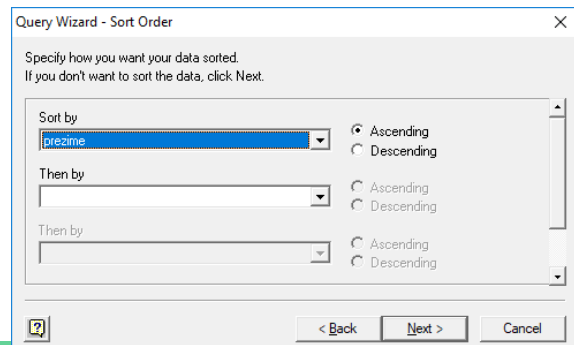
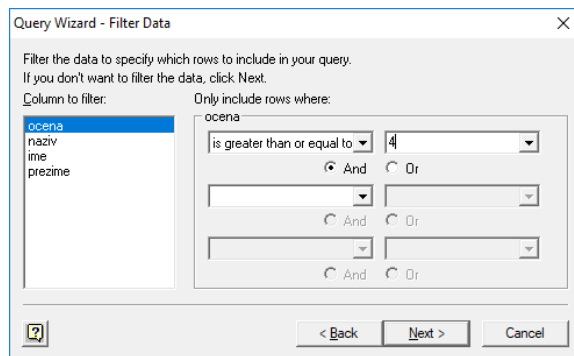
Uvoz podataka



114

114

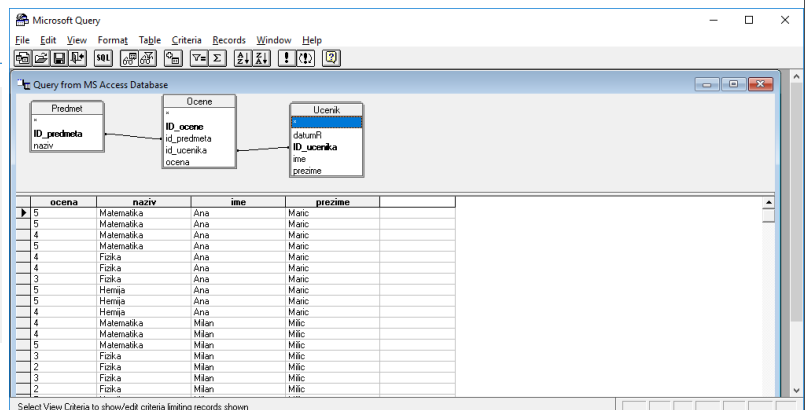
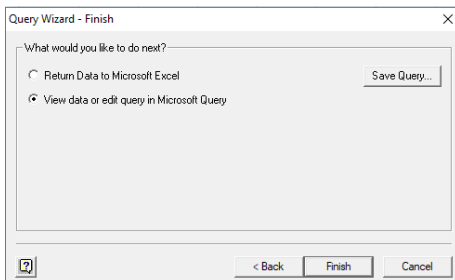
Uvoz podataka



115

115

Uvoz podataka



116

116

Uvoz podataka

	A	B	C	D
56	naziv	prezime	ime	ocena
57	Fizika	Maric	Ana	4
58	Fizika	Maric	Ana	4
59	Fizika	Maric	Ana	3
60	Fizika	Milic	Milan	3
61	Fizika	Milic	Milan	2
62	Fizika	Milic	Milan	3
63	Fizika	Milic	Milan	2
64	Fizika	Savic	Sanja	5
65	Fizika	Savic	Sanja	4
66	Hemija	Maric	Ana	5
67	Hemija	Maric	Ana	5
68	Hemija	Maric	Ana	4
69	Hemija	Milic	Milan	5
70	Hemija	Milic	Milan	5
71	Istorija	Milic	Milan	4
72	Istorija	Milic	Milan	4
73	Istorija	Milic	Milan	5
74	Matematika	Maric	Ana	5
75	Matematika	Maric	Ana	5
76	Matematika	Maric	Ana	4
77	Matematika	Maric	Ana	5
78	Matematika	Milic	Milan	4
79	Matematika	Milic	Milan	4

	A	B	C
1	Row Labels	Average of ocena	
2	Fizika	3,333333333	
3	Hemija	4,8	
4	Istorija	4,333333333	
5	Matematika	4,5	
6	Grand Total	4,148148148	
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			

PivotTable Field List

Choose fields to add to report:

ime

naziv

ocena

prezime

Drag fields between areas below:

Report Filter

Column Labels

Row Labels

Values

naziv

Average of o...

Defer Layout Update

Update

117

117

VBA (Visual Basic for Applications)

- VBA je programski jezik Microsoft Office programa (Excel, Word, Access)
 - pruža mogućnost automatizacije zadataka korišćenjem makroa (engl. *macros*)
 - proširuju se mogućnosti Office programa
 - Excel već pruža mnoštvo ugrađenih funkcija
 - za kompleksna izračunavanja je potrebno proširenje

118

118

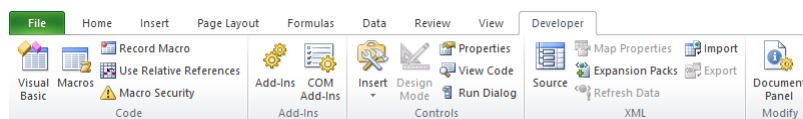
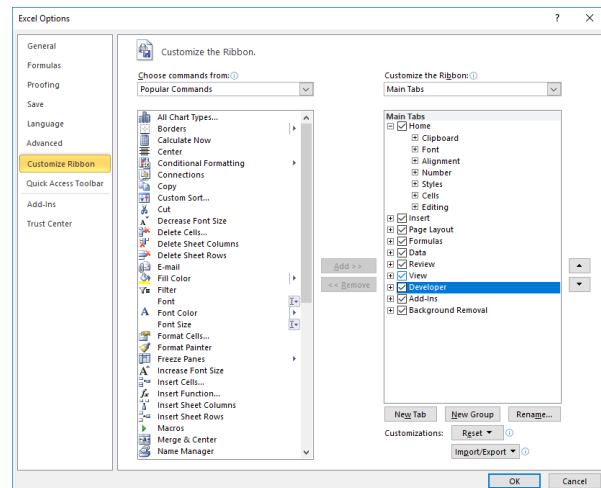
VBA (Visual Basic for Applications)

- Prvo je potrebno dodati meni za rad sa VBA
- Excel radni list koji sadrži makroe treba da bude snimljen kao Excel Enabled Macro Workbook (*.xlsm)

119

119

VBA (Visual Basic for Applications)



120

120

VBA (Visual Basic for Applications)

The screenshot illustrates the VBA development environment in Microsoft Excel. It shows the 'Developer' ribbon, the 'Form Controls' group, and the 'ActiveX Controls' group. A 'Hello' button is placed on the worksheet. A context menu is open over the button, with 'View Code' selected. The 'Properties' window shows the 'CommandButton1' object with various settings. The 'Microsoft Visual Basic for Applications' window shows the 'Private Sub Hello_Click()' procedure with the code 'MsgBox \"Hello World!\"'. Below the code window, a 'Hello World!' dialog box is displayed.

121

Range

- Range osobina predstavlja opseg ćelija u radnom listu
 - može biti jedna ćelija ili opseg ćelija
 - Range("A1")
 - Range("A1:A5")
 - Range("A1", "A5")
- Kada se obeleži opseg, na njega je moguće primeniti neku metodu
 - Range("A1").Select
 - Range("A2").Value = 10

122

122

Range

The screenshot displays the Microsoft Visual Basic for Applications environment. On the left, an Excel spreadsheet is visible with a range of cells (A14:B20) selected. A tooltip shows the command `Range("A14":"B20").Select`. The main window shows the VBA Project Explorer with the following structure:

- atpvbaen.xls (ATPVBAEN.XLAM)
 - Solver (SOLVER.XLAM)
 - VBAPROJECT (FUNCRES.XLAM)
 - VBAPROJECT (za slajdove.xlsx)
 - Microsoft Excel Objects
 - Sheet1 (Sheet1)
 - Sheet10 (kovarijansa)
 - Sheet11 (pivot)
 - Sheet12 (pivot 1)
 - Sheet13 (Sheet13)
 - Sheet14 (import iz txt)
 - Sheet15 (import access)
 - Sheet16 (import Query)
 - Sheet17 (import Query u pivot)
 - Sheet18 (Sheet9)

The VBA Code window on the right contains the following code:

```

Hello

Private Sub CommandButton1_Click()
    Range("A14:B20").Select
End Sub

```

123

123

Promenljive

- Promenljive služe za čuvanje podataka (brojevi, stringovi...) u memoriji
 - promenljive imaju ime i tip (Integer, String)
- U VBA potrebno je deklarirati promenljivu navođenjem ključne reči Dim
 - Dim a as Integer
- Nakon toga promenljivoj se dodeljuje vrednost
 - a = 10

124

124

Promenljive

23					
24					
25	10	20	Dodela Vrednosti U Celiju		
26	10	20			
27	10	20			
28	10	20			
29	10	20			
30	10	20			
31					
32					

```

Private Sub DodelaVrednostiUCeliju_Click()
    Dim a As Integer
    Dim b As Integer

    a = 10
    b = 20

    Range("A25:A30").Value = a
    Range("B25:B30").Value = b
End Sub

```

125

125

If naredba

- Uslovno izvršavanje naredbi


```

If Uslov Then
    naredbe
End If

```
- Uslov je logički izraz (True, False)


```

If Uslov Then
    naredbe
Else
    naredbe
End If

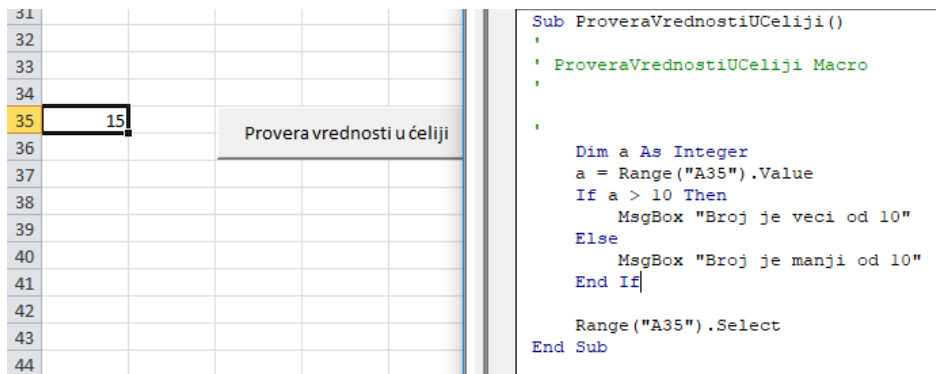
If Uslov Then
    naredbe
Elseif Uslov Then
    naredbe
Elseif Uslov Then
    naredbe
Else
    naredbe
End If

```

126

126

If naredba



31				
32				
33				
34				
35	15			
36				
37				
38				
39				
40				
41				
42				
43				
44				

```

Sub ProveraVrednostiUCeliji()
' ProveraVrednostiUCeliji Macro
'
'
Dim a As Integer
a = Range("A35").Value
If a > 10 Then
    MsgBox "Broj je veci od 10"
Else
    MsgBox "Broj je manji od 10"
End If
Range("A35").Select
End Sub

```

127

127

ActiveCell referenciranje

- Korišćenjem ActiveCell može se ukazati na neku drugu ćeliju pomoću brojeva vrsta i kolona
 - aktivna ćelija je u vrsti 1 i koloni 1 (početne „koordinate“)
 - pomeranje za jednu ćeliju u desno
 - ActiveCell(1,2)
 - pomeranje za jednu ćeliju u levo
 - ActiveCell(1,-1)
 - pomeranje za dve ćelije na dole
 - ActiveCell(3,1)

128

128

ActiveCell referenciranje

	A	B	C	D	E	F
1	Pera	91	10			
2	Mika	85	9			
3	Sava	95	10			
4	Maja	88	9			
5	Ana	100	10			
6	Luka	75	8			
7	Marko	35	5			
8	Maša	56	6			
9	Milan	75	8			
10	Dušan	89	9			
11	Ivana	93	10			
12	Sara	81	9			
13	Tanja	79	8			
14	Dejan	92	10			
15	Sanja	45	5			
16	Filip	99	10			
17						
18						
19						
20						

Ocene

```

Private Sub Ocene_Click()
    Dim bodovi As Integer
    bodovi = ActiveCell.Value
    If bodovi > 90 And bodovi <= 100 Then
        ActiveCell(1, 2).Value = 10
        ActiveCell(1, 2).Interior.Color = RGB(0, 255, 0)
    ElseIf bodovi > 80 And bodovi <= 90 Then
        ActiveCell(1, 2).Value = 9
        ActiveCell(1, 2).Interior.Color = RGB(0, 255, 200)
    ElseIf bodovi > 70 Then
        ActiveCell(1, 2).Value = 8
        ActiveCell(1, 2).Interior.Color = RGB(0, 255, 150)
    ElseIf bodovi > 60 Then
        ActiveCell(1, 2).Value = 7
        ActiveCell(1, 2).Interior.Color = RGB(0, 255, 100)
    ElseIf bodovi > 50 Then
        ActiveCell(1, 2).Value = 6
        ActiveCell(1, 2).Interior.Color = RGB(0, 255, 50)
    Else
        ActiveCell(1, 2).Value = 5
        ActiveCell(1, 2).Interior.Color = RGB(255, 0, 0)
    End If
End Sub

```

129

129

Select Case naredba

- Izbor iz liste vrednosti se može uraditi pomoću If-Elseif-Else naredbe

```

If izbor="R" Then
    MsgBox "Red"
Elseif izbor="G" Then
    MsgBox "Green"
Elseif izbor="B" Then
    MsgBox "Blue"
Else
    MsgBox "None"
EndIf

```

130

130

Select Case naredba

- Ali, moguće je koristiti i Select Case naredbu

Select Case Izbior

Case "R"

MsgBox "Red"

Case "G"

MsgBox "Green"

Case "B"

MsgBox "Blue"

Case Else

MsgBox "None"

End Select

131

131

Select Case naredba

24			
25			
26	Pera	91	10
27	Mika	85	9
28	Sava	95	10
29	Maja	88	9
30	Ana	100	10
31	Luka	75	8
32	Marko	35	5
33	Maša	56	6
34	Milan	75	8
35	Dušan	89	9
36	Ivana	93	10
37	Sara	81	9
38	Tanja	79	8
39	Dejan	92	10
40	Sanja	45	5
41	Filip	99	10
42			
43			
44			

Ocenjivanje

```
Private Sub Ocenjivanje_Click()
    Dim bodovi As Integer
    bodovi = ActiveCell(1, 1).Value
    Select Case bodovi
        Case 91 To 100
            ActiveCell(1, 2).Value = 10
            ActiveCell(1, 2).Interior.Color = RGB(0, 255, 0)
        Case 81 To 90
            ActiveCell(1, 2).Value = 9
            ActiveCell(1, 2).Interior.Color = RGB(0, 255, 200)
        Case 71 To 80
            ActiveCell(1, 2).Value = 8
            ActiveCell(1, 2).Interior.Color = RGB(0, 255, 150)
        Case 61 To 70
            ActiveCell(1, 2).Value = 7
            ActiveCell(1, 2).Interior.Color = RGB(0, 255, 100)
        Case 51 To 60
            ActiveCell(1, 2).Value = 6
            ActiveCell(1, 2).Interior.Color = RGB(0, 255, 50)
        Case Else
            ActiveCell(1, 2).Value = 5
            ActiveCell(1, 2).Interior.Color = RGB(255, 0, 0)
    End Select
End Sub
```

132

132

With

- Kada je potrebno pristupiti različitim osobinama aktivne ćelije
 - With object
 - .property
 - End With

133

133

With

44				
45				
46			Promena	
47				
48				
49				
50			Promena With	
51				
52				
53				
54				
55				
56				

```

Private Sub Promena_Click()
    ActiveCell.Font.Bold = True
    ActiveCell.Font.Color = vbBlue
    ActiveCell.Font.Name = "Arial"
    ActiveCell.Font.Size = 22
    ActiveCell.Font.Italic = True
End Sub

Private Sub PromenaWith_Click()
    With ActiveCell.Font
        .Bold = True
        .Color = vbBlue
        .Name = "Arial"
        .Size = 22
        .Italic = True
    End With
End Sub

```

134

134

For i For Each petlje

- For petlja se koristi za ponavljanje niza naredbi određen broj puta
 For a = 1 To 10
 MsgBox a
 Next a
- For Each petlja se koristi sa kolekcijama i nizovima
 For Each cell in Range("A1:A5")
 cell = Replace(cell.Value, " ", "_", "")
 Next cell

135

135

For i For Each petlje

	A	B	C	D	E	F	G	H	I	J	K
1		Matematika	Srpski	Fizika	Engleski	Informatika	Prosek				
2	Pera	5	5	5	5	5	5	odlican			
3	Mika	4	5	2	4	3	3,6	vrlo dobar			
4	Sanja	5	4	4	3	2	3,6	vrlo dobar			
5	Ana	5	3	5	4	4	4,2	vrlo dobar			
6	Milan	3	4	4	3	3	3,4	dobar			
7	Tanja	4	2	5	4	2	3,4	dobar			

```

Private Sub ForPetlja_Click()
    Dim prosek As Integer
    Dim i As Integer

    Dim broj As Integer
    broj = Range("A2:A20").Rows.Count

    For i = 2 To broj
        prosek = Range("G" & i).Value
        If prosek >= 4.5 Then
            Range("H" & i).Value = "odlican"
            Range("H" & i).Interior.Color = RGB(0, 255, 0)
        ElseIf prosek >= 3.5 Then
            Range("H" & i).Value = "vrlo dobar"
            Range("H" & i).Interior.Color = RGB(150, 255, 150)
        ElseIf prosek >= 2.5 Then
            Range("H" & i).Value = "dobar"
            Range("H" & i).Interior.Color = RGB(120, 255, 200)
        ElseIf prosek >= 1.5 Then
            Range("H" & i).Value = "dovoljan"
            Range("H" & i).Interior.Color = RGB(230, 255, 150)
        ElseIf prosek >= 1 Then
            Range("H" & i).Value = "nedovoljan"
            Range("H" & i).Interior.Color = RGB(240, 255, 200)
        End If
    Next i
End Sub

```

136

For i For Each petlje

53				
54				
55	Pera	91	10	Ocnjiv
56	Mika	85	9	
57	Sava	95	10	
58	Maja	88	9	
59	Ana	100	10	
60	Luka	75	8	
61	Marko	35	5	
62	Maša	56	6	
63	Milan	75	8	
64	Dušan	89	9	
65	Ivana	93	10	
66	Sara	81	9	
67	Tanja	79	8	
68	Dejan	92	10	
69	Sanja	45	5	
70	Filip	99	10	
71				
72				
73				
74				

```
Private Sub OcnjivanjeForEach_Click()
    Dim bodovi As Integer
    For Each cell In Range("B55:B70")
        bodovi = cell.Value
        Select Case bodovi
            Case 91 To 100
                cell(1, 2).Value = 10
                cell(1, 2).Interior.Color = RGB(0, 255, 0)
            Case 81 To 90
                cell(1, 2).Value = 9
                cell(1, 2).Interior.Color = RGB(0, 255, 200)
            Case 71 To 80
                cell(1, 2).Value = 8
                cell(1, 2).Interior.Color = RGB(0, 255, 150)
            Case 61 To 70
                cell(1, 2).Value = 7
                ActiveCell(1, 2).Interior.Color = RGB(0, 255, 100)
            Case 51 To 60
                cell(1, 2).Value = 6
                cell(1, 2).Interior.Color = RGB(0, 255, 50)
            Case Else
                cell(1, 2).Value = 5
                cell(1, 2).Interior.Color = RGB(255, 0, 0)
        End Select
    Next cell
End Sub
```

137

137

Do While petlja

- Do While petlja se koristi za ponavljanje niza naredbi, dokle god je zadovoljen uslov na početku

Do While uslov

naredbe

Loop

Dim brojac As Integer

brojac = 1

Do While brojac < 5

Cells(brojac, "D").Value = brojac

brojac = brojac + 1

Loop

138

138

Do While petlja

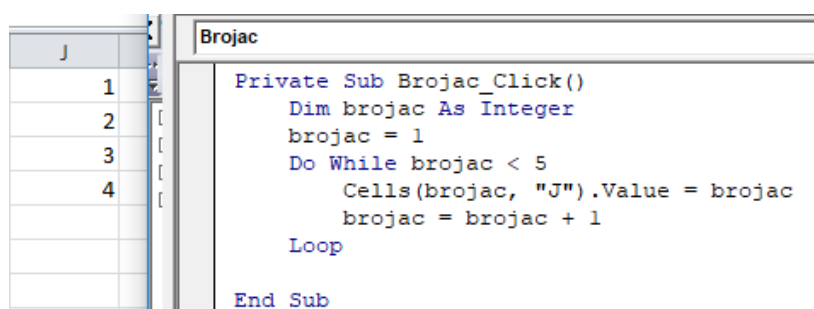
- Uslov je moguće postaviti i na kraju
Do
 naredbe
Loop While uslov
- Ovako se petlja sigurno izvršava bar jedan put

```
Dim brojac As Integer
brojac = 1
Do
    Cells(brojac, "D").Value = brojac
    brojac = brojac + 1
Loop While brojac < 5
```

139

139

Do While petlja



```
Private Sub Brojac_Click()
    Dim brojac As Integer
    brojac = 1
    Do While brojac < 5
        Cells(brojac, "J").Value = brojac
        brojac = brojac + 1
    Loop
End Sub
```

140

140

Do While petlja

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R		
1	Radnik	Radni sati					Ukupno	Plata												
2	Marko	6	8	10	9	8	41	20700		Private Sub IzracunajPlatu_Click()										
3	Milan	5	10	10	9	6	40	20000												
4	Ana	7	9	8	6	7	37	18500		Dim i As Integer										
5	Tanja	8	8	9	7	8	40	20000		Dim osnovna, prekovremeno, plata As Integer										
6	Filip	10	7	7	9	0	33	16500		For i = 2 To 12										
7	Saša	9	0	9	9	9	36	18000		If Range("G" & i).Value > 40 Then										
8	Sanja	6	9	8	8	6	37	18500		plata = (Range("G" & i).Value - 40) * Range("B17").Value + 40 * Range("B15").Value										
9	Maša	8	10	7	8	10	43	22100		Range("H" & i).Interior.Color = RGB(250, 0, 100)										
10	Igor	6	7	8	7	7	35	17500		Else										
11	Luka	9	9	8	10	6	42	21400		plata = Range("G" & i) * Range("B15")										
12	Dunja	7	8	7	6	7	35	17500		Range("H" & i).Interior.Color = RGB(0, 100, 259)										
13										End If										
14										Range("H" & i).Value = plata										
15	Cena sata	500								Next i										
16	Cena prekovremeno									End Sub										
17		700																		
18																				

141

141

Python i Pandas

142

Python

- Programski jezik
 - platformski nezavisan
 - objektno orijentisan
 - interpreterski
 - interaktivni
- Spada u grupu softvera otvorenog koda sa dobrom podrškom, dokumentacijom i pratećim bibliotekama

143

143

Python

- Podržava više paradigmi
 - proceduralno
 - funkcionalno
 - objektno-orijentisano
 - imperativno
- Akcenat je na efikasnosti programera i čitljivosti koda
- Nastao je u akademskim uslovima na univerzitetu Stičing u Holandiji
 - Guido van Rosum, 90. tih godina prošlog veka

144

144

Osnovne osobine jezika Python

- Nije strogo tipiziran jezik
- Blokovi se ne označavaju zagradama { i }, ili ključnim rečima *begin* i *end*
 - grupišu se uvlačenjem koda
- Komande nemaju znak ; na kraju
 - samo za komande koje su napisane u istom redu
- Dodela vrednosti predstavljena znakom =

145

145

Tipovi podataka

- Numerički
 - *int* – celi brojevi
 - *float* – brojevi u pokretnom zarezu
- String
 - sekvenca znakova između navodnika ("string", 'string')
 - pomoću indeksa pristupamo pojedinačnim znakovima
 - `s="string"`
 - `s[0]->s, s[-1]->g`
 - `s[1:3]->"tr"`

146

146

Kolekcije u Python-u

- `str` – string, nepromenljiva sekvenca karaktera
 - `s = "string"`
- `list` – promenljiva sekvenca objekata
 - `lista = [1, 2.3, "string", [1, 'a']]`
- `dict` – promenljivo mapiranje nepromenljivih ključeva na promenljive objekte
 - `rec = {"ime": "Ana", "prezime": "Peric"}`
- `tuple` – nepromenljiva sekvenca objekata
 - `torka = (1, "Novi Sad")`
- `range` – aritmetička progresija celobrojnih vrednosti
- `set` – promenljiva kolekcija jedinstvenih nepromenljivih objekata
 - `skup = {1, 2, 3}`

147

147

Python

- Paketi u Pythonu se nazivaju i moduli
- Paket se može posmatrati kao folder u kom su grupisane određene funkcije i tipovi podataka zajednički za neku oblast
- Da bi se mogle koristiti funkcije iz nekog paketa, moraju se prvo instalirati na lokalnom računaru i uvesti u program
 - `import <naziv_paketa>`
- Funkcija se poziva tako što se navodi ime paketa, pa onda ime funkcije

148

148

Python

- Poslednjih godina Python je poboljšao pakete koji se koriste za manipulaciju podacima, što ga stavlja u dobru poziciju za odabir jezika za analizu podataka
- Najčešće korišćeni paketi za analizu podataka u programskom jeziku Python su:
 - Pandas
 - Polars
 - NumPy
 - Matplotlib
 - ScyPy
 - Pytorch

149

149

Pandas

- Pandas (Panel Data) je biblioteka napisana u Python-u koja obezbeđuje manipulaciju podacima i analizu podataka
 - <https://pandas.pydata.org>
- Pandas obezbeđuje strukture podataka i funkcije projektovane tako da rad sa podacima bude brz i jednostavan

150

150

Pandas

- Primarni objekat u Pandas-u je okvir podataka (engl. *DataFrame*)
 - okvir podataka predstavlja **dvodimenzionalnu tabelarnu strukturu podataka**
- Omogućava čitanje i pisanje između memorije i fajlova različitih formata (json, xls, csv, ...)
- Prvobitno, Pandas je bio razvijen kao alat za aplikacije za analizu finansijskih podataka

151

151

Pandas

- Čitanje iz fajla i upisivanje u fajl
 - čitanje se obavlja pozivom Pandas funkcije i vraća se DataFrame objekat (`pandas.read_excel()`)
 - pisanje se obavlja pozivom funkcije nad DataFrame objektom (`DataFrame.to_csv()`)

Format Type	Data Description	Reader	Writer
text	CSV	read_csv	to_csv
text	JSON	read_json	to_json
text	HTML	read_html	to_html
text	Local clipboard	read_clipboard	to_clipboard
binary	MS Excel	read_excel	to_excel
binary	OpenDocument	read_excel	
SQL	SQL	read_sql	to_sql
SQL	Google Big Query	read_gbq	to_gbq

152

152

Pandas

- Za prikaz podataka mogu se koristiti razne funkcije
 - `head()`
 - `tail()`
 - `describe()`
- Za prikaz naziva kolona u DataFrame-u koristi se atribut `columns`
- Za sortiranje vrednosti po nekoj koloni koristi se funkcija `sort_values(by="naziv_kolone")`

153

153

Pandas

```
import pandas as pd
podaci = pd.read_csv("imena.csv")
print(podaci.describe())
print(podaci.head())
podaci.sort_values(by="prezime")
print(podaci.to_string())
```

```

                plata
count      11.000000
mean    44272.727273
std       6084.256883
min      34000.000000
25%      42000.000000
50%      45000.000000
75%      47500.000000
max       54000.000000
```

```

   ime prezime  plata
0    Ana   Peric  34000
1  Milan   Savic  35000
2  Dusan   Maric  43000
3  Tanja   Ilic  41000
4  Marko   Vasic  43000
5  Petar   Rakic  51000
6    Maja   Arsic  48000
7    Anja   Milic  45000
8    Sava   Popic  47000
9    Mila   Dakic  54000
10 Ivana   Ivic  46000
```

154

154

Pandas

- Selektovanje jedne kolone omogućeno je navođenjem imena kolone u uglastim zagradama
 - `df["naziv_kolone"]`
- Selektovanje jedne vrste moguće je upotrebom atributa `loc` i navođenjem indeksa
 - `df.loc[index]`
- Atribut `loc` može se koristiti i za selektovanje više kolona i vrsta
 - `df.loc[:, ["kolA", "kolB"]]`
- Takođe, može se koristiti i za selektovanje konkretne ćelije
 - `df.loc[5, "kolA"]`

155

155

Pandas

```
print(podaci["ime"])
```

0	Ana
1	Milan
2	Dusan
3	Tanja
4	Marko
5	Petar
6	Maja
7	Anja
8	Sava
9	Mila
10	Ivana

Name: ime, dtype: object

156

156

Pandas

```
print(podaci.loc[:, ["ime", "prezime"]])
```

```

           ime prezime
0      Ana    Peric
1    Milan   Savic
2    Dusan   Maric
3    Tanja   Ilic
4    Marko   Vasic
5    Petar   Rakic
6     Maja   Arsic
7     Anja   Milic
8     Sava   Popic
9     Mila   Dakic
10  Ivana   Ivic

```

```
print(podaci.loc[5, "prezime"])
```

```
Rakic
```

```
print(podaci.loc[5, ["ime", "prezime"]])
```

```

ime      Petar
prezime   Rakic
Name: 5, dtype: object

```

157

157

Pandas

- Moguće je filtrirati DataFrame i prikazati samo one elemente koji zadovoljavaju kriterijum
 - `df[df.kolona>5]`
- Za filtriranje se može koristiti i funkcija `isin()` koja proverava da li se u nizu nalazi tražena vrednost

158

158

Pandas

- Dodavanje nove kolone u DataFrame je jednostavno
 - definiše se ime kolone i dodeli niz vrednosti
 - `df["naziv_kolone"]=niz_vrednosti`
- Za postavljanje vrednosti u jednoj ćeliji koristi se funkcija `at()` i `iat()`
 - `df.at[v, "naziv_kolone"]=nova_vrednost`
 - `df.iat[v,k]=nova_vrednost`

159

159

Pandas

```
podaci["datumR"] = ["1.1.1990.", "2.2.1990.", "3.3.1990.", "4.4.1990.",
                   "5.5.1990.", "6.6.1990.", "7.7.1990.", "8.8.1990.", "9.9.1990.",
                   "10.10.1990.", "11.11.1990."]
```

```
print(podaci.to_string())
```

	ime	prezime	plata	datumR
0	Ana	Peric	34000	1.1.1990.
1	Milan	Savic	35000	2.2.1990.
2	Dusan	Maric	43000	3.3.1990.
3	Tanja	Ilic	41000	4.4.1990.
4	Marko	Vasic	43000	5.5.1990.
5	Petar	Rakic	51000	6.6.1990.
6	Maja	Arsic	48000	7.7.1990.
7	Anja	Milic	45000	8.8.1990.
8	Sava	Popic	47000	9.9.1990.
9	Mila	Dakic	54000	10.10.1990.
10	Ivana	Ivic	46000	11.11.1990.

160

160

Pandas

```
print (podaci["datumR"]>"7.7.1990.")
```

```
0    False
1    False
2    False
3    False
4    False
5    False
6    False
7     True
8     True
9    False
10   False
```

161

161

Pandas

- Pandas pri obradi ignoriše vrednosti koje nedostaju i koristi konstantu `nan` iz paketa NumPy kako bi to označio
- Za izbacivanje redova koje imaju `nan` vrednosti koristi se funkcija `dropna()`
- Za popunjavanje vrednosti koje nedostaju koristi se funkcija `fillna()`

162

162

Pandas

- Nad DataFrame-om moguće je vršiti grupisanje po nekom ključu
- Koristi se funkcija `groupby` i nakon toga se vrši neka agregacija nad grupisanim vrednostima
 - `df.groupby(["A", "B"]).sum()`
 - `df.groupby("A").sum()`

163

163

NumPy

- NumPy (Numerical Python) je osnovni paket za naučno računarstvo
 - www.numpy.org
- Sadrži posebnu strukturu višedimenzionalnih nizova (zamena za standardne liste u Python-u)
 - nizovi mogu sadržati samo elemente istog tipa
- Obezbeđuje:
 - operacije linearne algebre
 - statističke funkcije

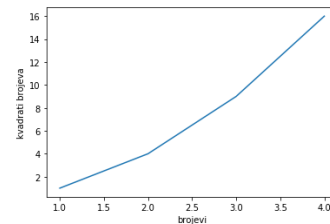
164

164

Matplotlib

- Matplotlib je najpopularnija biblioteka za izradu 2D grafika u Python-u
 - <https://matplotlib.org>
- Najčešće se koristi podmodul `pyplot`

```
import matplotlib.pyplot as plt
plt.plot([1, 2, 3, 4], [1, 4, 9, 16])
plt.xlabel('brojevi')
plt.ylabel('kvadrati brojeva')
plt.show()
```



165

165

ScyPy

- ScyPy je skup paketa koji se bave različitim standardnim problemima
- Neki od paketa su:
 - `scipy.integrate` – rešavanje diferencijalnih jednačina
 - `scipy.optimize` – optimizacija funkcija
 - `scipy.stats` – statistika

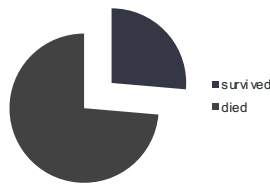
166

166

Primer - Excel

- Podaci o putnicima sa broda Titanic
<http://biostat.mc.vanderbilt.edu/wiki/pub/Main/DataSets/titanic3.xls>

pclass	survived	name	sex	age	sibsp	parch	ticket	fare	cabin	embarked	boat	body	home.dest
1	1	Allen, Miss. Elisabeth Walton	female	29	0	0	24160	211.3375	B5	S	2		St Louis, MO
1	1	Allison, Master. Hudson Trevor	male	0.9167	1	2	113781	151.5500	C22 C26	S	11		Montreal, PQ / Chesterville, ON
1	0	Allison, Miss. Helen Loraine	female	2	1	2	113781	151.5500	C22 C26	S			Montreal, PQ / Chesterville, ON
1	0	Allison, Mr. Hudson Joshua Creighton	male	30	1	2	113781	151.5500	C22 C26	S		135	Montreal, PQ / Chesterville, ON
1	0	Allison, Mrs. Hudson J C (Bessie Waklo Daniels)	female	25	1	2	113781	151.5500	C22 C26	S			Montreal, PQ / Chesterville, ON
1	1	Anderson, Mr. Harry	male	48	0	0	19952	26.5500	E12	S	5		New York, NY
1	1	Andrews, Miss. Komelia Theodosia	female	63	1	0	13502	77.9583	D7	S	10		Hudson, NY
1	0	Andrews, Mr. Thomas Jr	male	39	0	0	112050	0.0000	A36	S			Belfast, NI



167

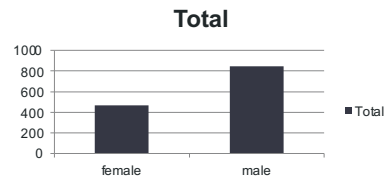
167

Primer - Excel

- Pivot tabela u kojoj je broj putnika po klasi putovanja
- Pivot tabela u kojoj je broj putnika prema polu

pclass	Total
1	323
2	277
3	709
Grand Total	1309

sex	Total
female	466
male	843
Grand Total	1309



168

168

Primer - Python

```
import pandas as pd
titanic = pd.read_excel("titanic3.xls")
print(titanic.describe())
```

	pclass	survived	...	fare	body
count	1309.000000	1309.000000	...	1308.000000	121.000000
mean	2.294882	0.381971	...	33.295479	160.809917
std	0.837836	0.486055	...	51.758668	97.696922
min	1.000000	0.000000	...	0.000000	1.000000
25%	2.000000	0.000000	...	7.895800	72.000000
50%	3.000000	0.000000	...	14.454200	155.000000
75%	3.000000	1.000000	...	31.275000	256.000000
max	3.000000	1.000000	...	512.329200	328.000000

```
[8 rows x 7 columns]
```

169

169

Primer - Python

```
print(titanic.info())
```

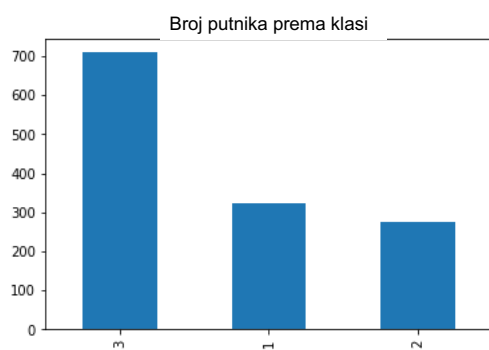
```
RangeIndex: 1309 entries, 0 to 1308
Data columns (total 14 columns):
pclass      1309 non-null int64
survived    1309 non-null int64
name        1309 non-null object
sex         1309 non-null object
age         1046 non-null float64
sibsp      1309 non-null int64
parch      1309 non-null int64
ticket     1309 non-null object
fare        1308 non-null float64
cabin       295 non-null object
embarked    1307 non-null object
boat        486 non-null object
body        121 non-null float64
home.dest   745 non-null object
dtypes: float64(3), int64(4), object(7)
memory usage: 143.2+ KB
None
```

170

170

Primer - Python

```
titanic['pclass'].value_counts().plot(kind='bar', title='Broj putnika prema klasi')
```

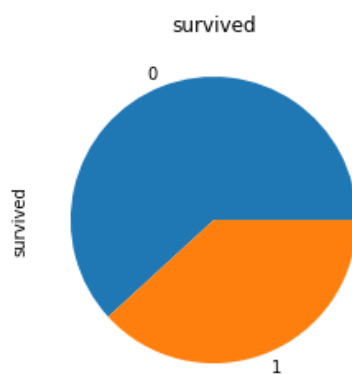


171

171

Primer - Python

```
titanic['survived'].value_counts().plot(kind='pie', title='survived')
```



172

172

Primer - Python

```
print("Najmlađi", titanic['age'].min())  
print("Najstariji", titanic['age'].max())  
print("Srednja vrednost", titanic['age'].mean())
```

```
Najmlađi 0.1667  
Najstariji 80.0  
Srednja vrednost 29.8811345124283
```

173

173

Primer - Dataset

- <https://www.opengov-muenchen.de/tr/dataset>
- <https://towardsdatascience.com/oktoberfest-quick-analysis-using-pandas-matplotlib-and-plotly-79778367a4c>
- Datase sadrži 8 kolona: godinu, trajanje, ukupan broj posetilaca, broj posetilaca po danu, cena piva, konzumacija piva, cena piletine, konzumacije piletine
- Zadatak
 - Istražiti kretanje cene piletine i piva na Oktoberfestu

174

174

Primer - Dataset

```
import pandas as pd
import matplotlib.pyplot as plt
df_oktoberfest = pd.read_csv('oktoberfestgesamt19852018.csv')
df_oktoberfest.head()
```

	jahr	dauer	besucher_gesamt	besucher_tag	bier_preis	bier_konsum	hendl_preis	hendl_konsum
0	1985	16	7.1	444	3.20	54541	4.77	629520
1	1986	16	6.7	419	3.30	53807	3.92	698137
2	1987	16	6.5	406	3.37	51842	3.98	732859
3	1988	16	5.7	356	3.45	50951	4.19	720139
4	1989	16	6.2	388	3.60	51241	4.22	775674

175

175

Primer - Dataset

- Kako se menjala cena piva u poslednje 33 godine? Kako se menjala cena piletine? Da li postoji pravilo?

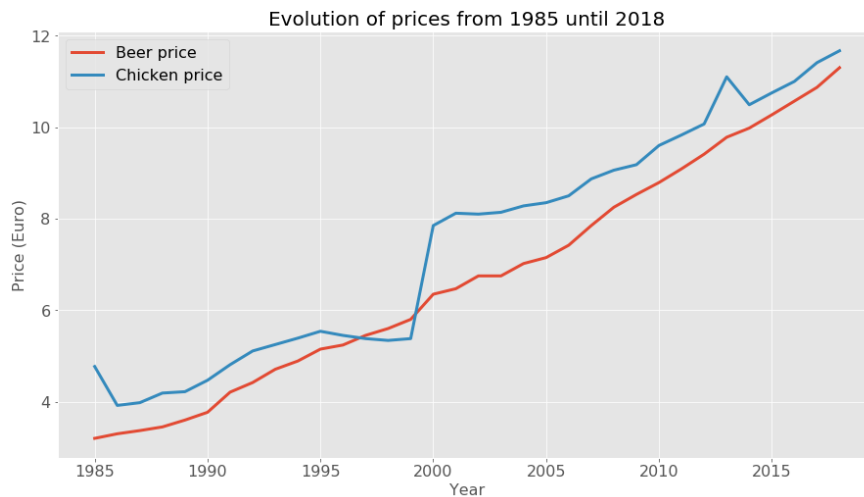
```
df_oktoberfest.plot(kind='line',
y=['beer_price', 'chicken_price'], figsize=(15, 8),
linewidth=3, fontsize=16)

plt.legend(labels=['Beer price', 'Chicken price'], fontsize=16)
plt.xlabel('Year', fontsize=16)
plt.ylabel('Price (Euro)', fontsize=16)
plt.title('Evolution of prices from 1985 until
2018', fontsize=20)
```

176

176

Primer - Dataset



177

177

Primer - Dataset

```
# Price of a beer and a chicken in 1985
print(df_oktoberfest.loc[1985,['beer_price','chicken_price']].sum())
# 7.97

# Price of a beer and a chicken in 2018
print(df_oktoberfest.loc[2018,['beer_price','chicken_price']].sum())
# 22.97
```

178

178

Primer - Dataset

- Kako se menjala potrošnja piva i piletine od 1985. do 2018. godine? Da li postoji pravilo?

```
df_oktoberfest.plot(y='beer_consumption', figsize=(15,8), linewidth=3, fontsize=16, color='maroon', label='_nolegend_')

years_landwirtschaftsfest = [1987, 1990, 1993, 1996, 2000, 2004, 2008, 2012, 2016]

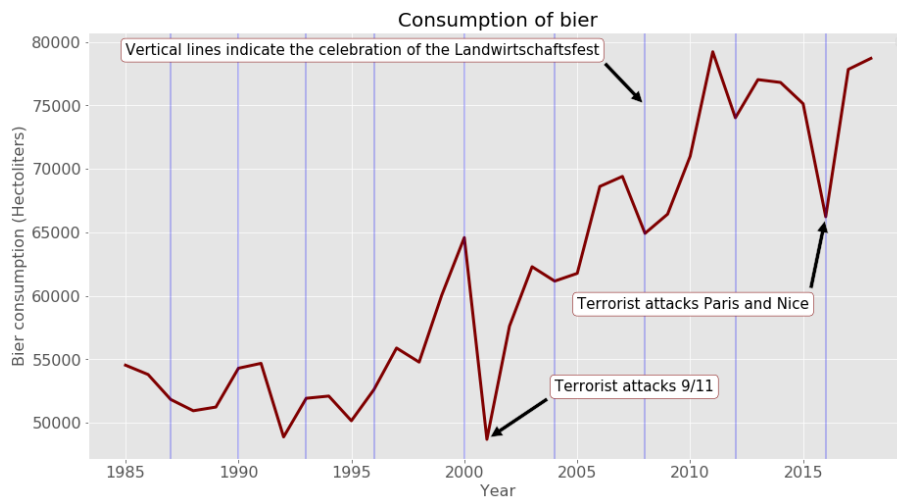
for year in years_landwirtschaftsfest:
    plt.axvline(x=year, color='blue', alpha=0.3)

plt.annotate('Vertical lines indicate the celebration of the Landwirtschaftsfest', xy=(2008, 75000), xytext=(1985, 79000),
            fontsize=15, arrowprops=dict(facecolor='black', shrink=0.05), bbox = dict(boxstyle="round", fc="w", ec="maroon"),)
plt.annotate('Terrorist attacks 9/11', xy=(2001, df_oktoberfest.loc[2001,'beer_consumption']), xytext=(2004, 52500),
            fontsize=15, arrowprops=dict(facecolor='black', shrink=0.05), bbox = dict(boxstyle="round", fc="w", ec="maroon"),)
plt.annotate('Terrorist attacks Paris and Nice', xy=(2016, df_oktoberfest.loc[2016,'beer_consumption']), xytext=(2005, 59000),
            fontsize=15, arrowprops=dict(facecolor='black', shrink=0.05), bbox = dict(boxstyle="round", fc="w", ec="maroon"),)
plt.xlabel('Year', fontsize=16)
plt.ylabel('Bier consumption (Hectoliters)', fontsize=16)
plt.title('Consumption of bier', fontsize=20)
```

179

179

Primer - Dataset



180

180

Primer - Dataset

```
df_oktoberfest.visitors_total.sort_values().head(3)
```

```

year
2001    5.5
2016    5.6
1988    5.7
Name: visitors_total, dtype: float64

```

181

181

Primer - Dataset

- Konzumacija piletine beleži pad

```
df_oktoberfest.plot(y='chicken_consumption', figsize=(15,8), linewidth=3, fontsize=16, color='darkblue',
label='_nolegend_')
```

```
years_landwirtschaftsfest = [1987, 1990, 1993, 1996, 2000, 2004, 2008, 2012, 2016]
```

```
for year in years_landwirtschaftsfest:
    plt.axvline(x=year, color='blue', alpha=0.3)
```

```
plt.annotate('Terrorist attacks 9/11', xy=(2001, df_oktoberfest.loc[2001,'chicken_consumption']),
xytext=(1991, 400000), fontsize=15, arrowprops=dict(facecolor='black', shrink=0.05),bbox =
dict(boxstyle="round", fc="w",ec="darkblue"),)
```

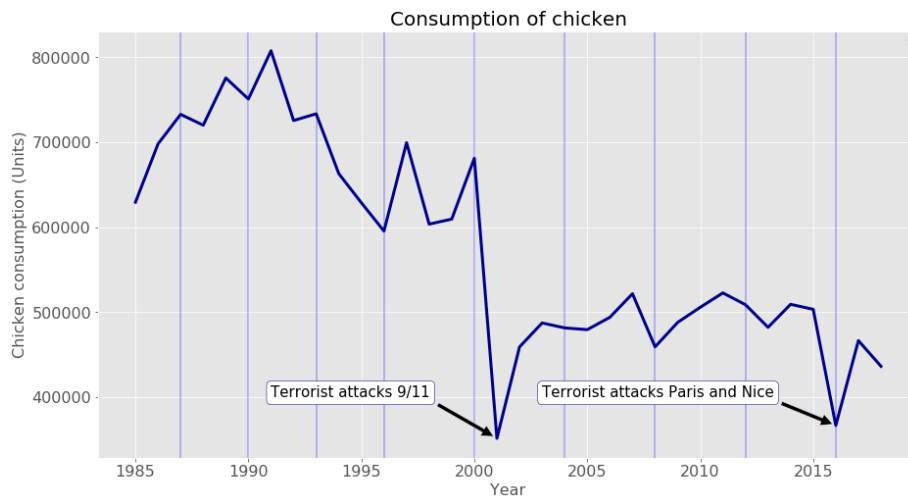
```
plt.annotate('Terrorist attacks Paris and Nice', xy=(2016,
df_oktoberfest.loc[2016,'chicken_consumption']), xytext=(2003, 400000), fontsize=15,
arrowprops=dict(facecolor='black', shrink=0.05),bbox = dict(boxstyle="round", fc="w",ec="darkblue"),)
```

```
plt.xlabel('Year',fontsize=16)
plt.ylabel('Chicken consumption (Units)',fontsize=16)
plt.title('Consumption of chicken',fontsize=20)
```

182

182

Primer - Dataset



183

183

Primer - Dataset

- Kako se menjao broj posetilaca od 1985. do 2018. godine?

```
df_oktoberfest.plot(y='visitors_total', figsize=(15,8), marker='o',
linewidth=2, fontsize=16, color='darkblue', label='_nolegend_')
```

```
plt.xlabel('Year',fontsize=16)
plt.ylabel('Number of visitors (Milions)',fontsize=16)
plt.title('Total number of visitors',fontsize=20)
```

184

184

Primer - Dataset

- Poštošnja piva se nije povećala zbog povećanog broja posetilaca već zato što ljudi više piju pivo



185

185

Primer - Dataset

- Rast cena prati linearni trend (posebno kod cena piva), pa se može primeniti linearna regresija da bi se predvidela cena piva i piletine 2030. godine

```

from sklearn.linear_model import LinearRegression

# Prediction of beer price
# Create a LinearRegression object
beer_regression = LinearRegression()
# Fit the model
beer_regression.fit(X=np.array(df_oktoberfest.index).reshape(-1, 1), y=df_oktoberfest.beer_price)
# Prediction using the linear model
print("Predicted beer price in 2030: {:.3f}€".format(beer_regression.predict(np.array([[2030]]))[0]))
# Predicted beer price in 2030: 13.781€

# Prediction of chicken price
# Create a LinearRegression object
chicken_regression = LinearRegression()
# Fit the model
chicken_regression.fit(X=np.array(df_oktoberfest.index).reshape(-1, 1), y=df_oktoberfest.chicken_price)
# Prediction using the linear model
print("Predicted chicken price in 2030: {:.3f}€".format(chicken_regression.predict(np.array([[2030]]))[0]))
# Predicted chicken price in 2030: 14.550€

```

2030. godine cena piva+piletine na
Oktoberfestu biće 28.33 €

186

186

Primer - Dataset

- Cene piva i piletine na Oktoberfestu rastu svake godine u poslednjih više od 30 godine
- Posetioci više piju, a manje jedu
 - u godinama terorističkih napada 2001. i 2016. godine, manje ljudi je posetilo festival
- Ukupan broj posetilaca ne prati nikakav trend, kreće se između 5,5 i 7,1 milion

187

187

R

- Programski jezik R je jezik i okruženje za statističko računarstvo i grafove
 - R je integrisani paket softverskih sredstava za manipulaciju podacima, proračun i grafičke prikaze
- Karakteristike programskog jezika R su:
 - besplatan je i otvorenog koda
 - radi na svim većim platformama: Windows, Linux, Mac
 - iste skripte se mogu koristiti na svim platformama
 - postoji velika, rastuća i aktivna zajednica korisnika R i kao rezultat toga, postoje brojni resursi za učenje i postavljanje pitanja
 - drugima je lako da doprinesu dodacima koji omogućavaju programerima da dele softverske implementacije novih metodologija

188

188

R - primer

- Rang lista ženske teniske asocijacije

```
library("xlsx", lib.loc="~/R/win-library/3.5")
wta<-read.xlsx("WTA.xlsx",sheetName="wta_ranking")
```

	first_name	last_name	country	height	service_speed	prize_money	singles_titles	w_matches	l_matches	hand
1	Naomi	Osaka	Japan	180	150	9635284	3	189	123	RIGHT
2	Simona	Halep	Romania	168	130	28955687	18	466	203	RIGHT
3	Petra	Kvitová	Czech Republic	182	170	19222229	22	500	264	LEFT
4	Karolína	Plíšková	Czech Republic	190	189	8391051	12	496	227	RIGHT
5	Angelique	Kerber	Germany	173	190	6184352	12	510	169	LEFT
6	Elina	Svitolna	Ukraine	174	169	6987556	13	320	96	RIGHT
7	Kiki	Bertens	Netherlands	182	192	8668863	8	333	142	RIGHT
8	Sloane	Stephens	United States	148	165	946396	4	200	81	RIGHT
9	Aisling	Bartny	Australia	166	111	1454822	5	199	93	RIGHT
10	Jaryn	Sabalenka	Belarus	182	190	609332	3	142	66	RIGHT
11	Serena	Williams	United States	177	194	16200000	72	759	155	RIGHT
12	Karoline	Wozniacki	Denmark	177	188	19668911	29	506	231	RIGHT
13	Anastasija	Svitolna	Latvia	169	169	5213602	3	299	162	RIGHT
14	Madison	Keys	United States	180	182	4587922	4	250	122	RIGHT
15	Qingye	Wang	China	190	199	7821859	9	99	33	RIGHT
16	Garbie	Muguruza	Spain	182	163	541236	7	320	111	RIGHT
17	Daria	Kasatkina	Russia	176	133	7589905	2	175	92	RIGHT
18	Donna	Vekic	Croatia	187	189	7890677	2	152	169	RIGHT
19	Lesia	Tsurenko	Ukraine	174	169	1105280	2	133	44	RIGHT

189

189

R - primer

- `str(wta)` - struktura podataka

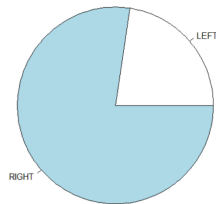
```
> str(wta)
'data.frame': 1800 obs. of 10 variables:
 $ first_name : Factor w/ 144 levels "Abbie","Aleksandra",...: 103 124 111 76 8 43 78 126 13 12 ...
 $ last_name  : Factor w/ 150 levels "Ailmer","Armsden",...: 97 54 72 104 66 124 12 120 7 115 ...
 $ country   : Factor w/ 19 levels "Australia","Belarus",...: 10 14 5 5 8 18 12 19 1 2 ...
 $ height    : num 180 168 182 190 173 174 182 148 166 182 ...
 $ service_speed : num 150 130 170 189 190 169 192 165 111 199 ...
 $ prize_money : num 9635284 28955687 19222229 8391051 6184352 ...
 $ singles_titles: num 3 18 22 12 12 13 8 4 5 3 ...
 $ w_matches  : num 189 466 500 496 510 320 333 200 199 142 ...
 $ l_matches  : num 123 203 264 227 169 96 142 81 93 66 ...
 $ hand      : Factor w/ 2 levels "LEFT","RIGHT": 2 2 1 2 1 2 2 2 2 2 ...
```

190

190

R - primer

- `pie(table(wta$hand))`



- Python:

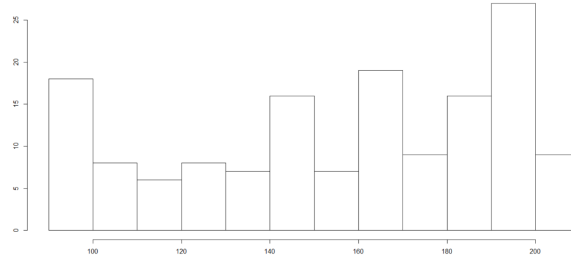
```
wta['hand'].value_counts().plot(kind='pie', title='Hand')
```

191

191

R - primer

- `hist(wta$service_speed)`



- Python:

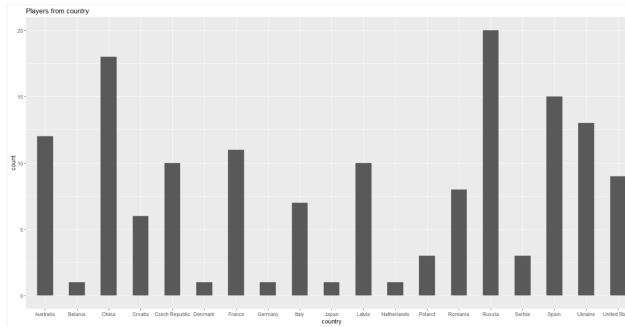
```
import pylab as p
p.hist(wta.service_speed)
```

192

192

R - primer

- `ggplot(wta, axes(x=country)) + geom_bar(width=0.5) + ggtitle("'Players from country'")`



- Python:

```
wta['country'].value_counts().plot(kind='bar', title='Players from country')
```

193

193

R - primer

- `min(wta$height)`
- `max(wta$height)`
- `mean(wta$height)`

```
> min(wta$height)
[1] 141
> max(wta$height)
[1] 202
> mean(wta$height)
[1] 176.4133
> |
```

- Python:

```
wta['height'].min()
wta['height'].max()
wta['height'].mean()
```

194

194

Literatura

- Dalje učenje
 - <https://kellyfoulk.herokuapp.com/post/3>

195