



## **Strukture šeme BP Data Warehouse sistema**

---

*Logički i dimenzioni model  
DW sistema*

# Sadržaj

- Projektovanje DW sistema
- Projektovanje šeme BP
- Zvezdasta šema
- Strukture činjeničnih podataka
- Strukture dimenzionih podataka
- Hijerarhije dimenzija
- Šema tipa pahuljice
- Šema tipa sazvežđa
- Šema sa agregiranim podacima

# Projektovanje DW sistema

---

- Tipične aktivnosti
  - analiza i specifikacija korisničkih zahteva
    - specifikacija poslovnog modela
  - projektovanje šeme DW BP
    - **projektovanje konceptualne šeme DW BP**
      - **specifikacija logičkog modela**
    - **projektovanje implementacione šeme DW BP**
      - **specifikacija dimenzionog modela**
    - projektovanje fizičke organizacije šeme DW BP
      - specifikacija fizičkog modela
  - projektovanje arhitekture DW sistema
  - projektovanje ECTL softverske podrške
  - projektovanje softverske podrške za izveštavanje i analizu podataka

# Sadržaj

---

- Projektovanje DW sistema
- Projektovanje šeme BP
- Zvezdasta šema
- Strukture činjeničnih podataka
- Strukture dimenzionih podataka
- Hijerarhije dimenzija
- Šema tipa pahuljice
- Šema tipa sazvežđa
- Šema sa agregiranim podacima

# Projektovanje šeme BP

---

- Projektovanje konceptualne šeme DW BP
  - specifikacija logičkog modela
    - bliskog korisnicima
    - nezavisnog od tehnološke platforme i ciljnog SUBP
    - visoko formalnog
      - mogućnost konceptualnog iskazivanja dimenzionih struktura
    - tehnike projektovanja – upotreba
      - ER modela podataka
      - UML jezika i modela (dijagrama) klasa
- Projektovanje implementacione šeme DW BP
  - specifikacija dimenzionog modela
    - generisanog transformacijom logičkog modela
    - pripremljenog za ciljni SUBP – danas najčešće RSUBP
    - upotreba relacionog modela podataka

# Sadržaj

---

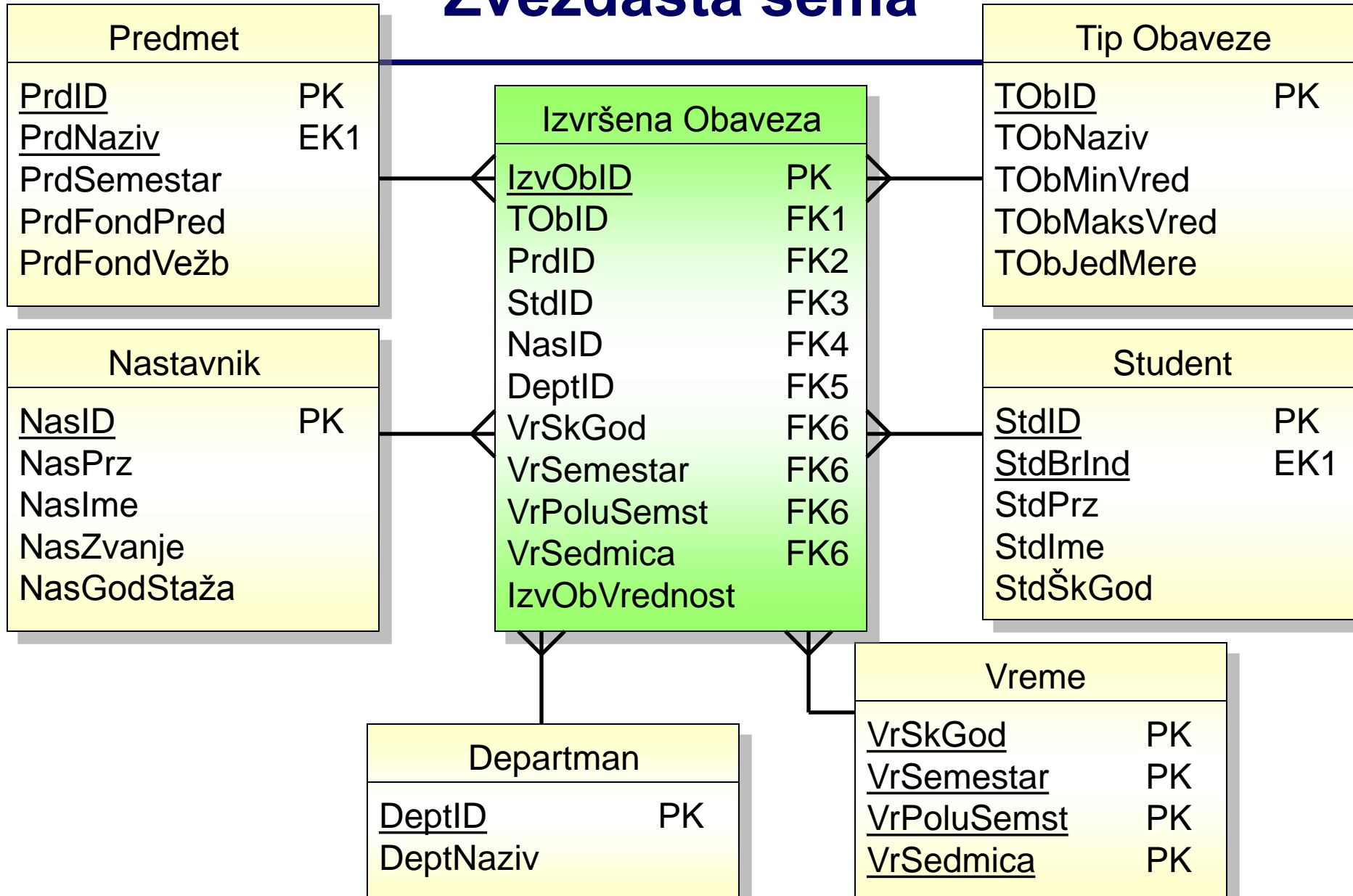
- Projektovanje DW sistema
- Projektovanje šeme BP
- Zvezdasta šema
- Strukture činjeničnih podataka
- Strukture dimenzionih podataka
- Hijerarhije dimenzija
- Šema tipa pahuljice
- Šema tipa sazvežđa
- Šema sa agregiranim podacima

# Zvezdasta šema

---

- **Star Schema**
- Komponente
  - **strukture činjeničnih podataka (činjenica)**
    - tipovi entiteta / šeme relacija činjenica
      - sa definisanim primarnim ključem
      - sa definisanim atributima i domenima mera - činjenica
  - **strukture dimenzionih podataka**
    - tipovi entiteta / šeme relacija dimenzija
      - sa definisanim skupovima ključeva i primarnim ključevima
      - sa definisanim atributima dimenzija
      - sa definisanim hijerarhijama dimenzija

# Zvezdasta šema





# Zvezdasta šema

---

- Osobine
  - lako razumljiva struktura
  - omogućava sprovođenje multidimenzionih analiza
  - podržana od strane velikog broja alata za izveštavanje i analizu podataka
- Uslovi normalizacije
  - očekivano zadovoljava uslov 1NF
  - šeme relacija činjenica najčešće zadovoljavaju BCNF
  - šeme relacija dimenzija mogu, ali često ne zadovoljavaju uslove >1NF

# Sadržaj

---

- Projektovanje DW sistema
- Projektovanje šeme BP
- Zvezdasta šema
- **Strukture činjeničnih podataka**
- Strukture dimenzionih podataka
- Hijerarhije dimenzija
- Šema tipa pahuljice
- Šema tipa sazvežđa
- Šema sa agregiranim podacima

# Strukture činjeničnih podataka

---

- **Šema relacije (tabela) činjenica**
  - naziva se i **tabela detaljnih podataka**
  - nalazi se uvek u centru zvezdaste šeme
  - referencira tabele dimenzija putem stranih ključeva koji predstavljaju primarne ključeve dimenzija
  - sadrži attribute mera
    - reprezentuju numeričke podatke – pokazatelje poslovanja
    - često reprezentuju agregirane (sumarne) podatke
  - relacija može da sadrži ekstremno veliki broj torki
  - količina podataka može da raste brzo u vremenu
  - tipično aditivna
    - relacija se ažurira najčešće upisivanjem novih torki

# Strukture činjeničnih podataka

---

- **Ključevi šeme relacije činjenica**
  - uobičajeno sadrži jedan ključ – primarni ključ
  - tipična struktura primarnog ključa
    - **veštački ključ (surrogate key)**
    - negovoreća šifra
    - jedan numerički atribut
    - nova vrednost se generiše inkrementiranjem prethodne
  - dozvoljena struktura primarnog ključa
    - složeni ključ
    - dobijen uniranjem svih prenetih primarnih ključeva tabela dimenzija
    - ređe u primeni
      - lošije performanse, nego u slučaju upotrebe veštačkog ključa

# Strukture činjeničnih podataka

---

- **Vrste tabela činjenica**
  - **sa uključenim merama**
    - **Regular Fact Table**
    - postoje atributi koji reprezentuju mere – vrednosti činjenica
  - **bez uključenih mera**
    - **Factless Fact Table**
    - ne postoje atributi koji reprezentuju vrednosti činjenica

# Strukture činjeničnih podataka

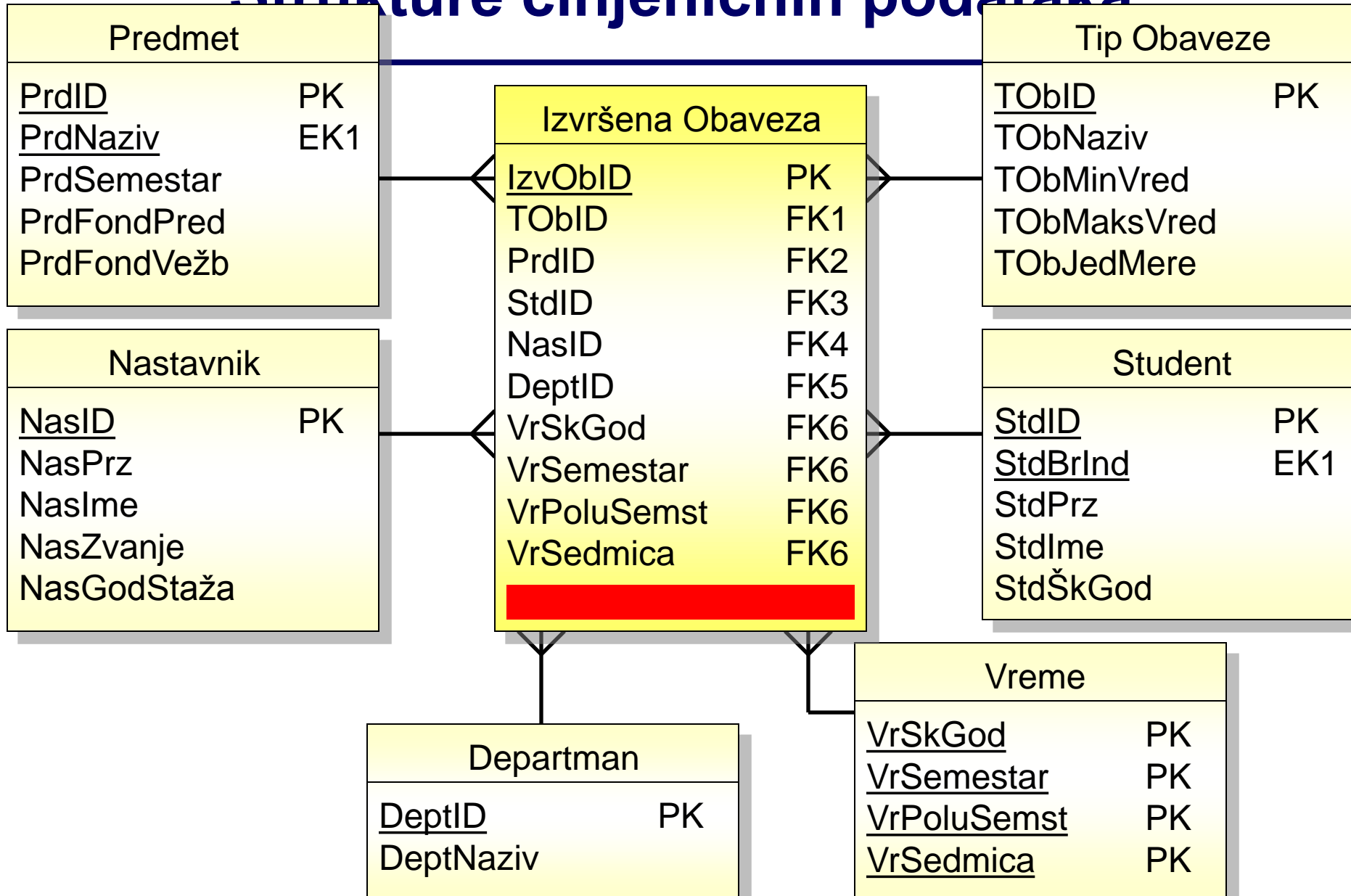
---

- **Factless Fact Table**

- uobičajeno služi za praćenje događaja
  - event tracking
- dva moguća stanja za bilo koju kombinaciju vrednosti dimenzija
  - postoji toraka u relaciji činjenica – događaj se zbio
  - ne postoji toraka u relaciji činjenica – događaj se nije zbio
- često reprezentuje strukturu tzv. retkih matrica
- upiti sa agregacijom podataka tipično koriste funkciju COUNT
- Primer
  - praćenje posećenosti nastave ili izvršenja obaveza na predmetu od strane studenta, do nivoa sedmice ili npr. pojedinačnog održanog časa



# Strukture činjeničnih podataka



# Strukture činjeničnih podataka

---

- **Vrste mera činjenica**
  - **prema mogućnosti sumiranja po dimenzijama**
    - pri realizaciji upita nad izabranim dimenzijama
  - **aditivne mere činjenica**
    - **Additive Fact Measures**
    - numeričke vrednosti, sabirljive po svim dimenzijama
  - **semiaditivne mere činjenica**
    - **Semiadditive Fact Measures**
    - numeričke vrednosti, sabirljive samo po nekim dimenzijama
  - **neaditivne mere činjenica**
    - **Nonadditive Fact Measures**
    - numeričke ili nenumeričke vrednosti, nesabirljive po dimenzijama
    - mogu biti osnova za generisanje drugih, sabirljivih dimenzija



# Strukture činjeničnih podataka

- Primeri
  - dimenzije: *kupac, region, vreme, proizvod*

Mera činjenice	Vrsta	Komentar
<i>naplaćeni iznos</i>		
<i>prodata količina</i>		
<i>ocena zadovoljstva kupca, na skali [1..5]</i>		
činjenica bez uključenih mera (factless data)		

# Strukture činjeničnih podataka

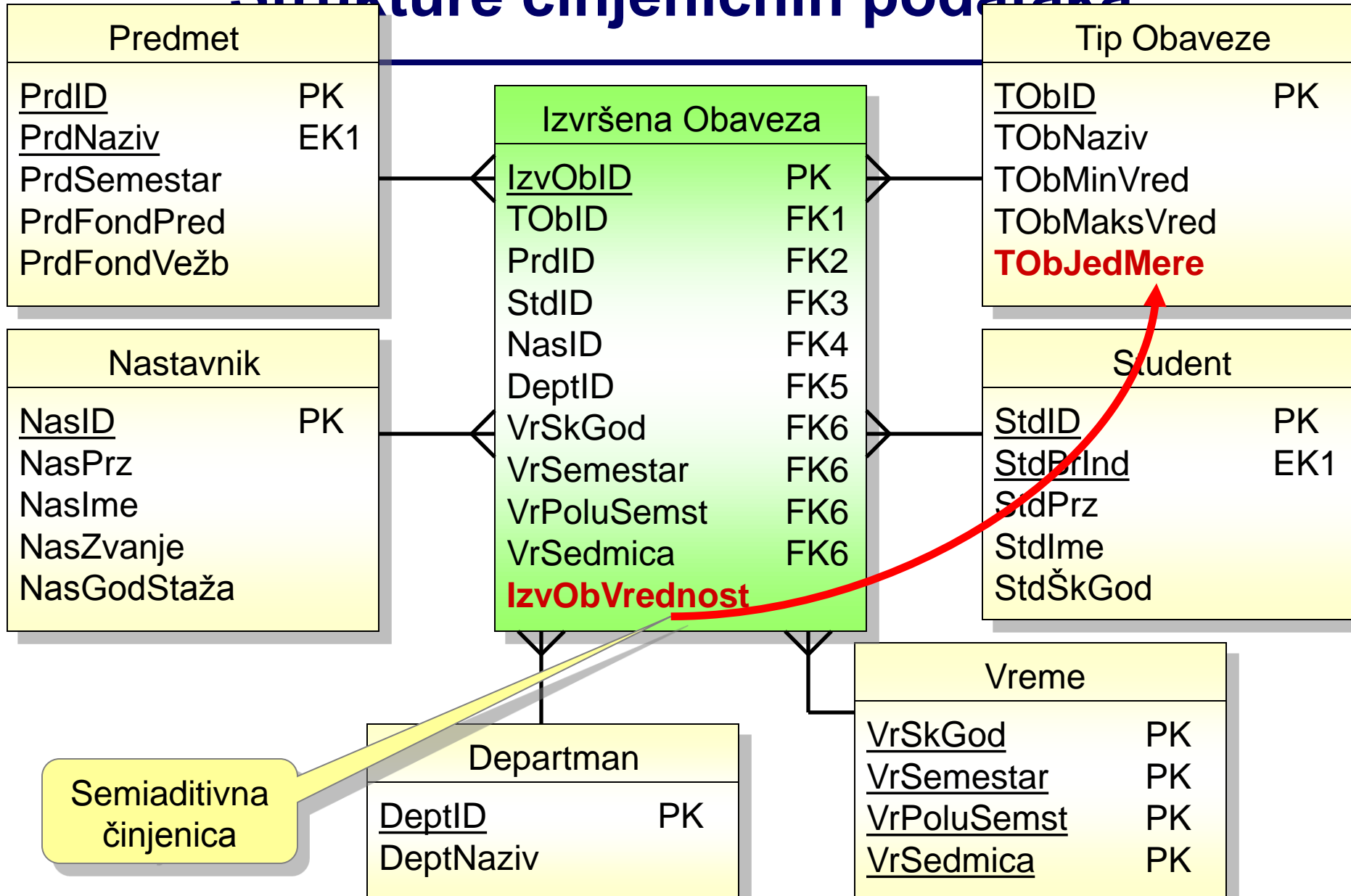
- Primeri

- dimenzije: *kupac, region, vreme, proizvod*

Mera činjenice	Vrsta	Komentar
<i>naplaćeni iznos</i>	<b>aditivna</b>	<ul style="list-style-type: none"> <li>• sabirljiva po bilo kojoj kojoj kombinaciji dimenzija</li> </ul>
<i>prodata količina</i>	<b>semiaditivna</b>	<ul style="list-style-type: none"> <li>• sabirljiva po proizvodu, odnosno</li> <li>• po bilo kojoj kombinaciji dimenzija koja uključuje specificirani proizvod</li> </ul>
<i>ocena zadovoljstva kupca, na skali [1..5]</i>	<b>neaditivna</b>	<ul style="list-style-type: none"> <li>• nesabirljiva po navedenim dimenzijama</li> <li>• dozvoljava izračunavanje npr. srednjih vrednosti i odstupanja</li> </ul>
činjenica bez uključenih mera (factless data)	<b>neaditivna</b>	<ul style="list-style-type: none"> <li>• nesabirljiva po navedenim dimenzijama</li> <li>• dozvoljava prebrojavanje pojava činjenica</li> </ul>



# Strukture činjeničnih podataka



# Strukture činjeničnih podataka

---

- **Vrste mera činjenica**
  - prema izvoru generisanja vrednosti
  - **osnovne mere činjenica**
    - **Base Fact Measures**
    - vrednosti atributa se generišu preuzimanjem podataka iz različitih izvora
  - **izvedene (izračunate) mere činjenica**
    - **Derived Fact Measures**
    - vrednosti atributa se izvode (izračunavaju) iz vrednosti drugih činjenica

# Strukture činjeničnih podataka

- Primeri
  - dimenzije: *kupac, region, vreme, proizvod*

Mera činjenice	Vrsta	Komentar
<i>naplaćeni iznos</i>		<ul style="list-style-type: none"> <li>• preuzimaju se i sumiraju podaci dnevnih transakcija iz OLTP BP</li> </ul>
<i>prodata količina</i>		<ul style="list-style-type: none"> <li>• preuzimaju se i sumiraju podaci dnevnih transakcija iz OLTP BP</li> </ul>
<i>efikasnost prodaje, na skali [1..3]</i>		<ul style="list-style-type: none"> <li>• izračunavanje vrednosti na nivou DSA, po osnovu ocenjivanja visine količnika:  <math display="block">K = (\text{naplaćeni iznos} / \text{prodata količina}) * 100</math> <ul style="list-style-type: none"> <li><math>K &lt; 10000 \Rightarrow \text{EfP} = 1</math></li> <li><math>K \text{ BETWEEN } 10000 \text{ AND } 20000 \Rightarrow \text{EfP} = 2</math></li> <li><math>K &gt; 20000 \Rightarrow \text{EfP} = 3</math></li> </ul> </li> </ul>

# Strukture činjeničnih podataka

- Primeri

- dimenzije: *kupac, region, vreme, proizvod*

Mera činjenice	Vrsta	Komentar
<i>naplaćeni iznos</i>	<b>osnovna</b>	• preuzimaju se i sumiraju podaci dnevnih transakcija iz OLTP BP
<i>prodata količina</i>	<b>osnovna</b>	• preuzimaju se i sumiraju podaci dnevnih transakcija iz OLTP BP
<i>efikasnost prodaje, na skali [1..3]</i>	<b>izvedena</b>	<ul style="list-style-type: none"> <li>• izračunavanje vrednosti na nivou DSA, po osnovu ocenjivanja visine količnika:  <math display="block">K = (\text{naplaćeni iznos} / \text{prodata količina}) * 100</math> <ul style="list-style-type: none"> <li><math>K &lt; 10000 \Rightarrow \text{EfP} = 1</math></li> <li><math>K \text{ BETWEEN } 10000 \text{ AND } 20000 \Rightarrow \text{EfP} = 2</math></li> <li><math>K &gt; 20000 \Rightarrow \text{EfP} = 3</math></li> </ul> </li> </ul>

# Sadržaj

---

- Projektovanje DW sistema
- Projektovanje šeme BP
- Zvezdasta šema
- Strukture činjeničnih podataka
- **Strukture dimenzionih podataka**
- Hijerarhije dimenzija
- Šema tipa pahuljice
- Šema tipa sazvežđa
- Šema sa agregiranim podacima

# Strukture dimenzionih podataka

---

- **Šeme relacija (tabele) dimenzija**
  - reprezentuju kontekst činjenica
  - povezane su sa šemama relacija činjenica
    - prostiranjem primarnog ključa u šeme relacija činjenica
- **sadrže attribute koji**
  - opisuju samu dimenziju (tekstualno, numerički, vizuelno) ili
  - reprezentuju agregirane (sumarne) numeričke podatke po dimenziji
- **relacija dimenzije**
  - može, po obimu podataka, biti i mala i velika
    - tipično, bitno manji obim podataka u odnosu na relacije činjenica
  - sadrži, po pravilu, relativno retko promenljive podatke



# Strukture činjeničnih podataka

---

- **Ključevi šeme relacije dimenzije**
  - može sadržati više ekvivalentnih ključeva
    - jedan izabrani – primarni ključ
  - izbor i struktura primarnog ključa
    - **veštački ključ**
      - formiran namenski za potrebe DW sistema ili organizacije
      - negovoreća šifra
      - jedan numerički atribut
      - nova vrednost se generiše inkrementiranjem prethodne
    - **prirodni ključ**
      - nasleđen iz poslovanja i/ili izvornog OLTP sistema
      - nekada složen iz više atributa – obimne vrednosti
        - » nepogodan u primeni
        - » jer se prenosi kao strani ključ u tabelu činjenica
        - » problem performansi pri realizaciji spajanja tabela

# Strukture činjeničnih podataka

---

- **Izbor ključa šeme relacije dimenzije**

- problem

- izbor granularnosti dimenzije – uticaj na izbor ključa

- primer: Praćenje prodaje automobila, **dimenzija:** proizvod

- dilema: Šta predstavlja dimenziju proizvoda?

- » Marka automobila?

- » Marka i model automobila?

- » Marka, model i varijanta ("paket opcija") automobila?

- » Šta je sa uključivanjem dodatnih opcija, po želji korisnika?

- posledice

- izbor granularnosti dimenzije – često i strateška odluka

- izbor nedovoljno fine granularnosti – jedan od uzroka propasti DW sistema u praksi

- izbor suviše fine granularnosti – "eksplozija" DW podataka

- savet: izabrati za jedan nivo finiju granularnost od potreba

# Strukture činjeničnih podataka

---

- **Izbor ključa šeme relacije dimenzije**

- problem

- **nejednaka jedinična dimenzija** u različitim izvorima podataka
  - primer: Praćenje prodaje automobila, **dimenzija:** proizvod
    - » u različitim OLTP sistemima, ključ proizvoda ne ukazuje na isti nivo granularnosti
    - » u jednom sistemu ključ ukazuje na marku i model, a u drugom na marku, model i varijantu

- posledice

- nemogućnost lakog "poravnavanja" granularnosti dimenzije u DW sistemu na željeni (finiji) nivo
  - problemi u formiranju ključa dimenzije proizvod
- dilema
  - da li izabrati grublji (mogući) nivo granularnosti, ili
  - pokušati sa interpolacijom podataka na finijem nivou

# Strukture činjeničnih podataka

---

- **Izbor ključa šeme relacije dimenzije**

- problem

- **upotreba različitih sistema označavanja** osnovnih činilaca u različitim izvorima podataka
  - različite vrednosti ključa za isti činilac poslovanja
  - primer: Praćenje prodaje automobila, **dimenzija**: proizvod
    - » u različitim OLTP sistemima, ključ proizvoda formira se na različite načine
    - » ključ proizvoda se ne mora sastojati ni iz istih atributa
    - » čak i kada se sastoji iz istih atributa, domeni i sistematika dodele vrednosti se razlikuju

- posledice

- pojava istih entiteta dimenzije s različitim vrednostima ključa
- pojava različitih entiteta dimenzije s istim vrednostima ključa
- dilema: Kako prepoznati entitete - homonime i sinonime?

# Strukture činjeničnih podataka

---

- **Izbor ključa šeme relacije dimenzije**

- problem

- **mogućnost modifikacije podataka dimenzije**

- iako su dimenzioni podaci relativno statički, nije isključena mogućnost njihovog ažuriranja

- "zastarevanje" dimenzije u vremenu

- posledice

- podela jedne dimenzije u vremenu na više drugih

- primer: podela poslovnog partnera na dva nova

- spajanje više dimenzija u jednu

- primer: spajanje dva poslovna partnera u jednog

- modifikacija klasifikacionih podataka dimenzije

- primer: domaći poslovni partner postaje inostrani (promena granica države)

- savet: ublažiti problem izborom finije granularnosti dimenzije

# Strukture činjeničnih podataka

---

- **Izmenljive dimenzije - operacije**
  - **upis nove torke dimenzije**
    - bez ograničenja, ako to nema uticaja na ostale torke dimenzije
  - **brisanje postojeće torke dimenzije**
    - neprimereno, ako postoje povezane vrednosti činjenica
  - **izmena podataka jedne torke dimenzije**
    - pristupi
      - ne vodi se evidencija o istoriji izmena podataka
      - vodi se evidencija o istoriji izmena podataka
  - **podela dimenzije u vremenu na više drugih**
    - zahteva evidenciju istorije promena dimenzija
  - **spajanje više dimenzija u jednu**
    - zahteva evidenciju istorije promena dimenzija

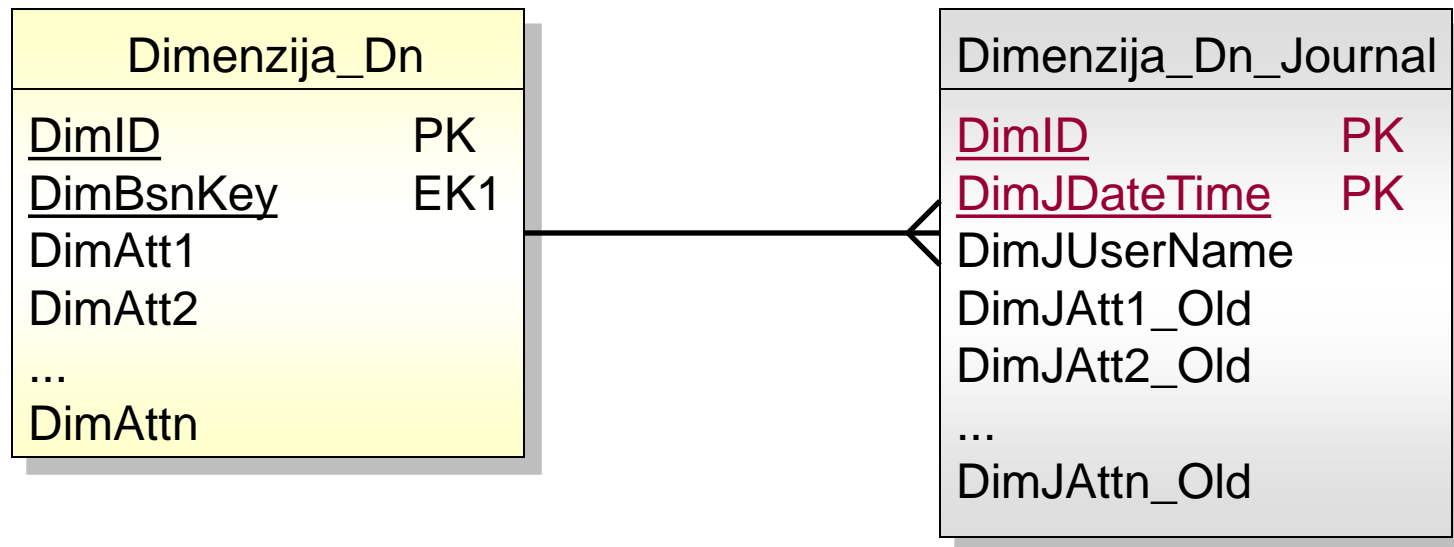
# Strukture činjeničnih podataka

---

- **Izmena podataka jedne torke dimenzije**
  - bez evidencije istorije izmena podataka
    - **tehnika "overwrite"**
      - prepisivanje novih podataka preko starih
      - nema uticaja na vrednost primarnog ključa torke dimenzije i povezane vrednosti činjenica
      - ne postoji evidentirana istorija promena torke dimenzije
      - ne postoje evidentirani podaci o izvršiocu promena
  - vodi se evidencija o istoriji izmena podataka
    - tehnike
      - **"istorijski podaci u posebnoj tabeli"**
      - **"istorijski podaci u istoj tabeli"**
      - **"istorijski podaci u posebnim kolonama"**

# Strukture činjeničnih podataka

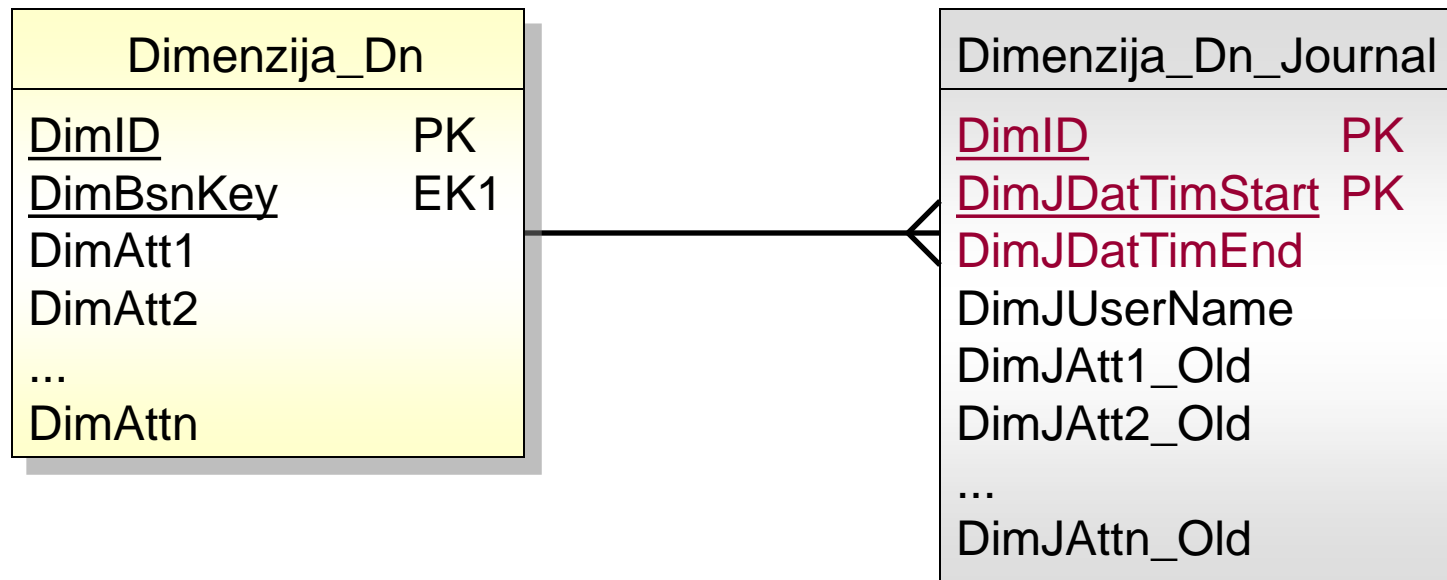
- **Izmena podataka jedne torke dimenzije**
  - "istorijski podaci u posebnoj tabeli" (1)
    - torka se dodaje u istorijsku tabelu pri svakom upisu i modifikaciji, ili samo pri modifikaciji torke, tabele dimenzije
    - pogodno za podršku podela i spajanja dimenzija
      - ako se u tabelu dimenzije uvede binarni atribut DimAkt "aktuelna dimenzija",  $dom(DimAkt) = \{ 'Y', 'N' \}$





# Strukture činjeničnih podataka

- **Izmena podataka jedne torke dimenzije**
  - "istorijski podaci u posebnoj tabeli" (2)
    - torka se dodaje u istorijsku tabelu pri svakom upisu i modifikaciji, ili samo pri modifikaciji torke, tabele dimenzije
    - pogodno za podršku podela i spajanja dimenzija
      - može se u tabelu dimenzije uvesti binarni atribut DimAkt "aktuelna dimenzija",  $dom(DimAkt) = \{ 'Y', 'N' \}$



# Strukture činjeničnih podataka

- **Izmena podataka jedne torke dimenzije**
  - "istorijski podaci u istoj tabeli" (1)
    - nova torka sa istim DimID se upisuje, umesto modifikacije atributa DimAtt1, ..., DimAttn, originalne torke
    - nepogodno za podršku podela i spajanja dimenzija
      - podeljene / spojene torke morale bi imati drugačiji DimID

Dimenzija_Dn	
<u>DimID</u>	PK
<u>DimDateTime</u>	PK
DimBsnKey	
DimAtt1	
DimAtt2	
...	
DimAttn	
DimUserName	

Dimenzija_Dn	
<u>DimID</u>	PK
<u>DimDatTimStart</u>	PK
DimDatTimEnd	
DimBsnKey	
DimAtt1	
DimAtt2	
...	
DimAttn	
DimUserName	

# Strukture činjeničnih podataka

- **Izmena podataka jedne torke dimenzije**
  - "istorijski podaci u istoj tabeli" (2)
    - nova torka sa istim DimID se upisuje, umesto modifikacije atributa DimAtt1, ..., DimAttn, originalne torke
    - pogodno za direktno evidentiranje činjenica po verzijama date dimenzije
    - pogodno za podršku podela i spajanja dimenzija
      - ako se vrednosti za DimVersionNo hijerarhijski organizuju

Dimenzija_Dn	
<u>DimID</u>	PK
<u>DimVersionNo</u>	PK
DimDatTimStart	
DimDatTimEnd	
DimBsnKey	
DimAtt1	
DimAtt2	
...	
DimAttn	
DimUserName	

# Strukture činjeničnih podataka

- **Izmena podataka jedne torke dimenzije**
  - "istorijski podaci u posebnim kolonama" (1)
    - modifikuje se torka prepisivanjem starih vrednosti u posebne kolone (DimAtt1\_Old,..., DimAttn\_Old)
    - evidentira se samo prethodno i tekuće stanje
    - nepogodno za podršku podela i spajanja dimenzija

Dimenzija_Dn	
<u>DimID</u>	PK
<u>DimBsnKey</u>	EK1
DimAtt1	
...	
DimAttn	
DimAtt1_Old	
...	
DimAttn_Old	
DimDateTime	
DimUserName	

# Strukture činjeničnih podataka

- **Intervalna (klasifikaciona) dimenzija**

- **"Bracketing Dimension"**

- definiše klasifikaciju činjeničnih podataka po jednom ili više kriterijuma

- definisanih putem (intervalnih) vrednosti dimenzionih atributa

- primer

Tipovi korisnika	
<u>DInID</u>	PK
DInPol	
DInBračSt	
DInStDob	
DInPrihod	

DInID	DInPol	DInBračSt	DInStDob	DInPrihod
1	Muški	Sam	< 25	< 10K
2	Muški	Sam	< 25	[10K..20K]
3	Muški	Sam	< 25	> 20K
4	Muški	Sam	[25..34]	< 10K
...	...	...	...	...
47	Ženski	U braku	> 55	[10K..20K]
48	Ženski	U braku	> 55	> 20K

# Sadržaj

---

- Projektovanje DW sistema
- Projektovanje šeme BP
- Zvezdasta šema
- Strukture činjeničnih podataka
- Strukture dimenzionih podataka
- Hijerarhije dimenzija
- Šema tipa pahuljice
- Šema tipa sazvežđa
- Šema sa agregiranim podacima

# Hijerarhije dimenzija

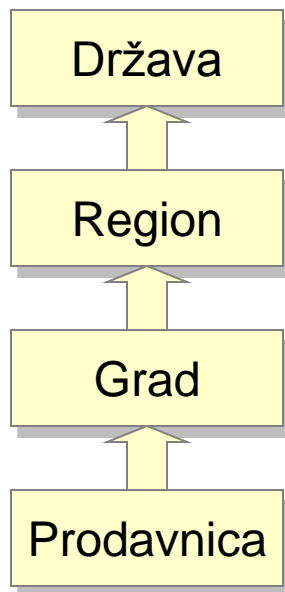
- **Hijerarhija dimenzije**

- reprezentuje višenivoovsku, hijerarhijsku strukturu dimenzionih podataka
- obezbeđuje višenivoovsku klasifikaciju činjeničnih podataka po datoj dimenziji
- atributi dimenzije mogu se organizovati u jednu ili više hijerarhijskih struktura
  - odnos između tipova entiteta po nivoima hijerarhije:  $N : 1$
  - definisan nizom funkcionalnih zavisnosti između atributa dimenzije
  - primer
    - *ProdavnicaID* → *GradID*,
    - *GradID* → *RegionID*,
    - *RegionID* → *DržavaID*

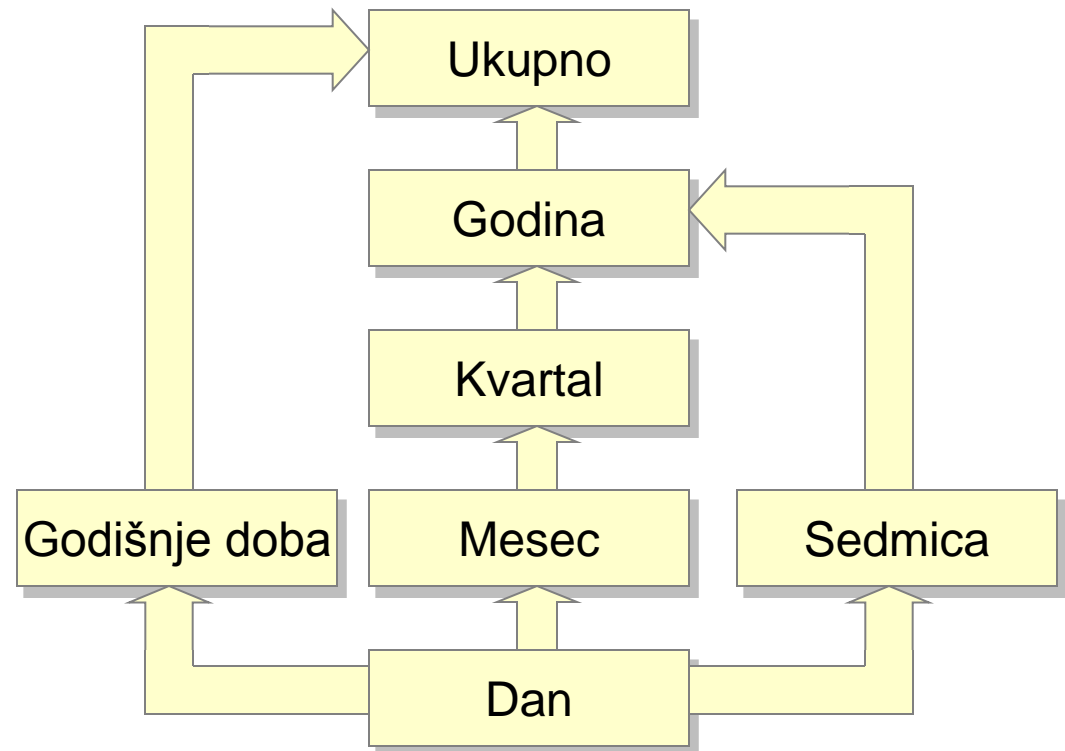
# Hijerarhije dimenzija

- Primeri

(A) Hijerarhija prostorne dimenzije



(B) Višestruka hijerarhija vremenske dimenzije





# Hijerarhije dimenzija

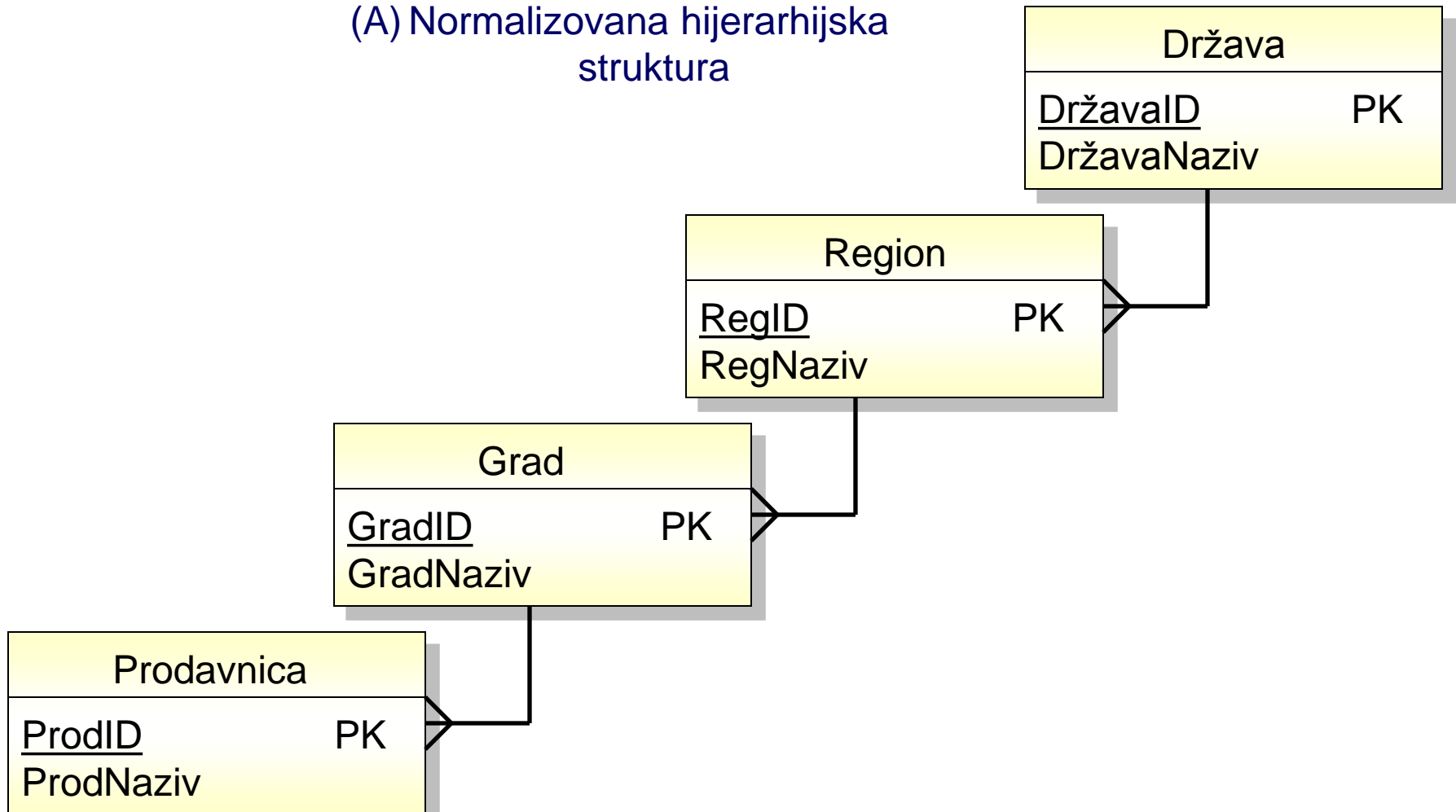
---

- **Hijerarhija dimenzije**
  - načini modelovanja u DW šemi BP
    - **(A) normalizovana struktura**
      - jedan hijerarhijski nivo = jedna šema relacije
    - **(B) delimično denormalizovana struktura**
      - više hijerarhijskih nivoa = jedna šema relacije
    - **(C) kompletno denormalizovana struktura**
      - svi hijerarhijski nivoi = jedna šema relacije

# Hijerarhije dimenzija

- Primeri

(A) Normalizovana hijerarhijska struktura



# Hijerarhije dimenzija

- Primeri

(C) Kompletno denormalizovana  
hijerarhijska struktura

Prostorna struktura	
<u>ProdID</u>	PK
ProdNaziv	
GradID	
GradNaziv	
RegID	
RegNaziv	
DržavaID	
DržavaNaziv	

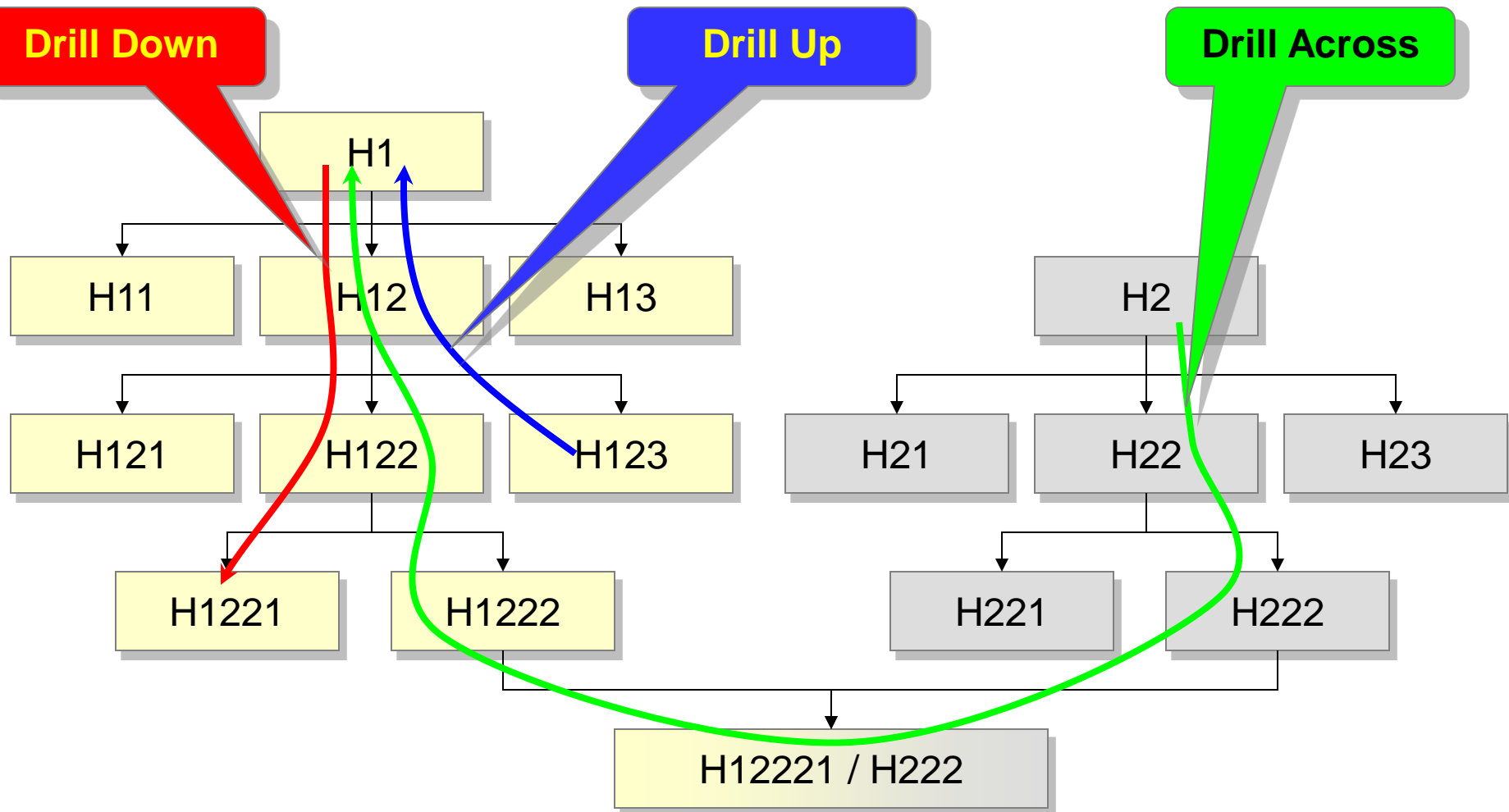
# Hijerarhije dimenzija

---

- **Upotreba hijerarhija dimenzija**
  - za analizu činjenica na različitim nivoima agregacije
    - od viših ka nižim nivoima
      - **Drill Down**
    - od nižih ka višim nivoima
      - **Drill Up (Roll Up)**
  - za analizu činjenica po različitim hijerarhijama
    - kombinacija viši – niži nivo (hijerarhija A) – niži – viši nivo (hijerarhija B)
      - **Drill Across**
  - za obezbeđenje optimizacije upita, zasnovane na prestrukturiranju specifikacije upita
    - **Query Rewriting**

# Hijerarhije dimenzija

- Upotreba hijerarhija dimenzija - ilustracija



# Hijerarhije dimenzija

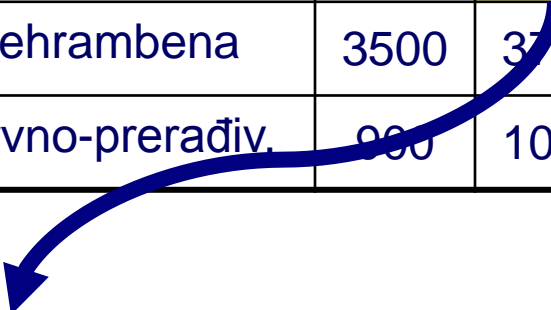
---

- Primer – Drill Down
  - zapažanje
    - signifikantna razlika godišnjeg pokazatelja poslovanja za određenu granu industrije
  - drill down analiza podataka po hijerarhiji dimenzije
    - pronaći proizvode koji imaju najznačajniji uticaj na pad vrednosti posmatranog pokazatelja

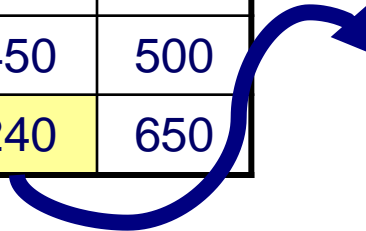
# Hijerarhije dimenzija

- Primer – Drill Down

Industrija	2000	2001	2002
Automobilska	1500	940	1480
Prehrambena	3500	3750	3650
Drvno-prerađiv.	900	1000	950



Automobilska	2000	2001	2002
Kamionska	300	250	330
Autobuska	500	450	500
Porodič. auto	700	240	650



Porodič. auto	2000	2001	2002
Escort	500	40	370
Falcon	100	100	130
Mustang	100	100	150

# Sadržaj

---

- Projektovanje DW sistema
- Projektovanje šeme BP
- Zvezdasta šema
- Strukture činjeničnih podataka
- Strukture dimenzionih podataka
- Hijerarhije dimenzija
- Šema tipa pahuljice
- Šema tipa sazvežđa
- Šema sa agregiranim podacima



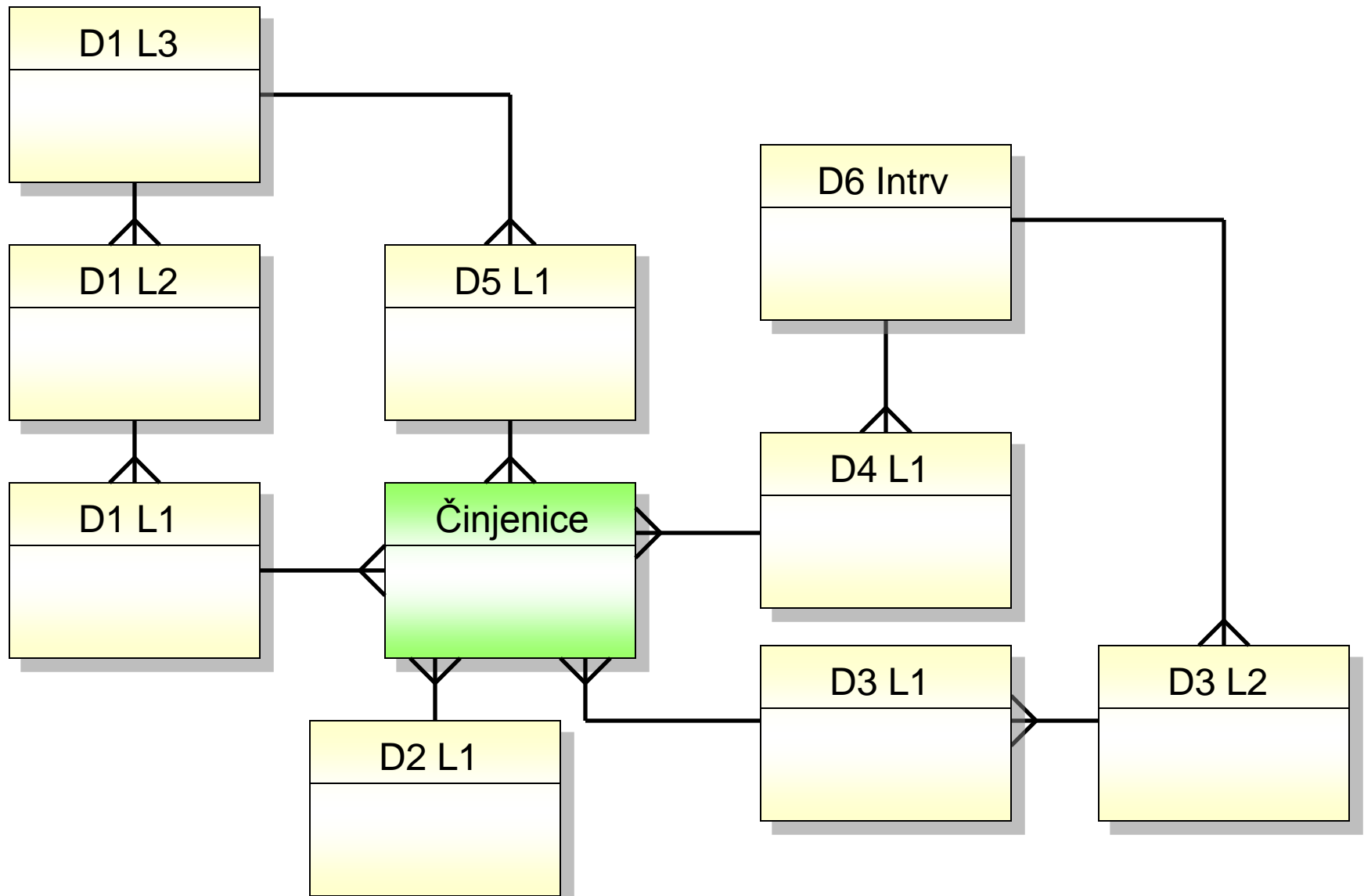
# Šema tipa pahuljice

---

- **Snowflake Schema**

- varijanta zvezdaste šeme
- hijerarhije dimenzija modelovane putem normalizovane strukture
  - šeme relacija dimenzija zadovoljavaju uslov BCNF
  - povezane su putem prostiranja primarnog ključa
- prednosti
  - izbegavanje logičkih problema koji su posledica narušavanja BCNF/3NF
  - struktura eksplicitno iskazuje hijerarhije dimenzija i moguće nivoe agregacije
- nedostaci
  - problem performansi upita, kao posledica potrebe spajanja torki iz različitih relacija dimenzija

# Šema tipa pahuljice



# Sadržaj

---

- Projektovanje DW sistema
- Projektovanje šeme BP
- Zvezdasta šema
- Strukture činjeničnih podataka
- Strukture dimenzionih podataka
- Hijerarhije dimenzija
- Šema tipa pahuljice
- Šema tipa sazvežđa
- Šema sa agregiranim podacima

# Šema tipa sazvežđa

---

- **Constellation Schema**
  - kombinacija više zvezdastih šema koje dele zajedničke dimenzije
  - motivacija
    - DW često sadrži više tabela činjenica
      - u cilju podrške različitih tema poslovanja
    - postoje dimenzije koje su zajedničke za različite tabele činjenica

# Šema tipa sazvežda

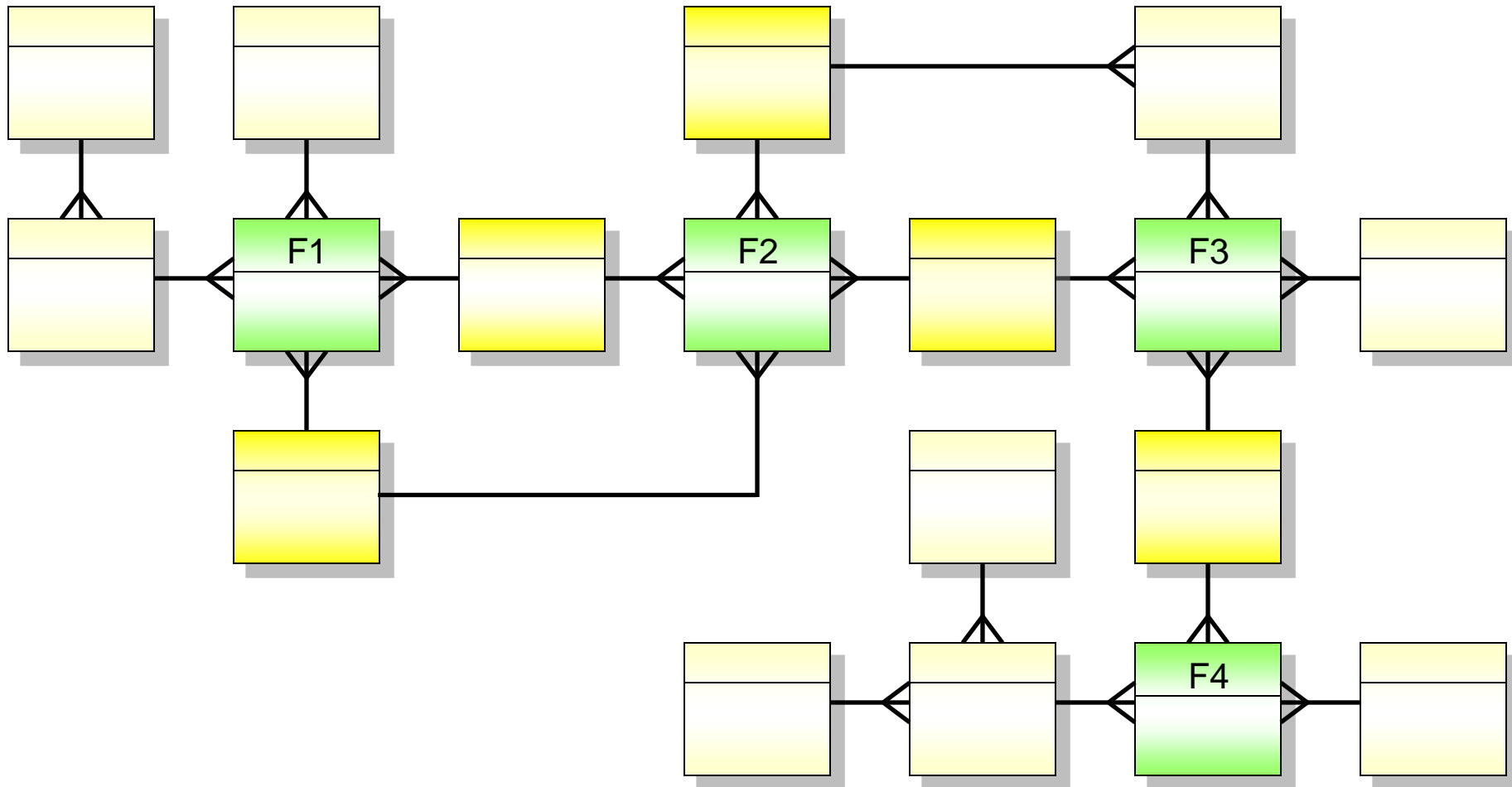
- **Constellation Schema**

– primer

Dimenzija	Tema / Proces poslovanja	
	Efikasnost stud.	Prihodi od stud.
<b>Vreme</b>	×	×
Nastavnik	×	
<b>Studijski prog.</b>	×	×
Predmet	×	
<b>Student</b>	×	×
Vrste prihoda		×

- Deljene dimenzije za navedene tabele činjenica
  - *Vreme*, *Studijski program* i *Student*

# Šema tipa sazvežđa



# Sadržaj

---

- Projektovanje DW sistema
- Projektovanje šeme BP
- Zvezdasta šema
- Strukture činjeničnih podataka
- Strukture dimenzionih podataka
- Hijerarhije dimenzija
- Šema tipa pahuljice
- Šema tipa sazvežđa
- Šema sa agregiranim podacima

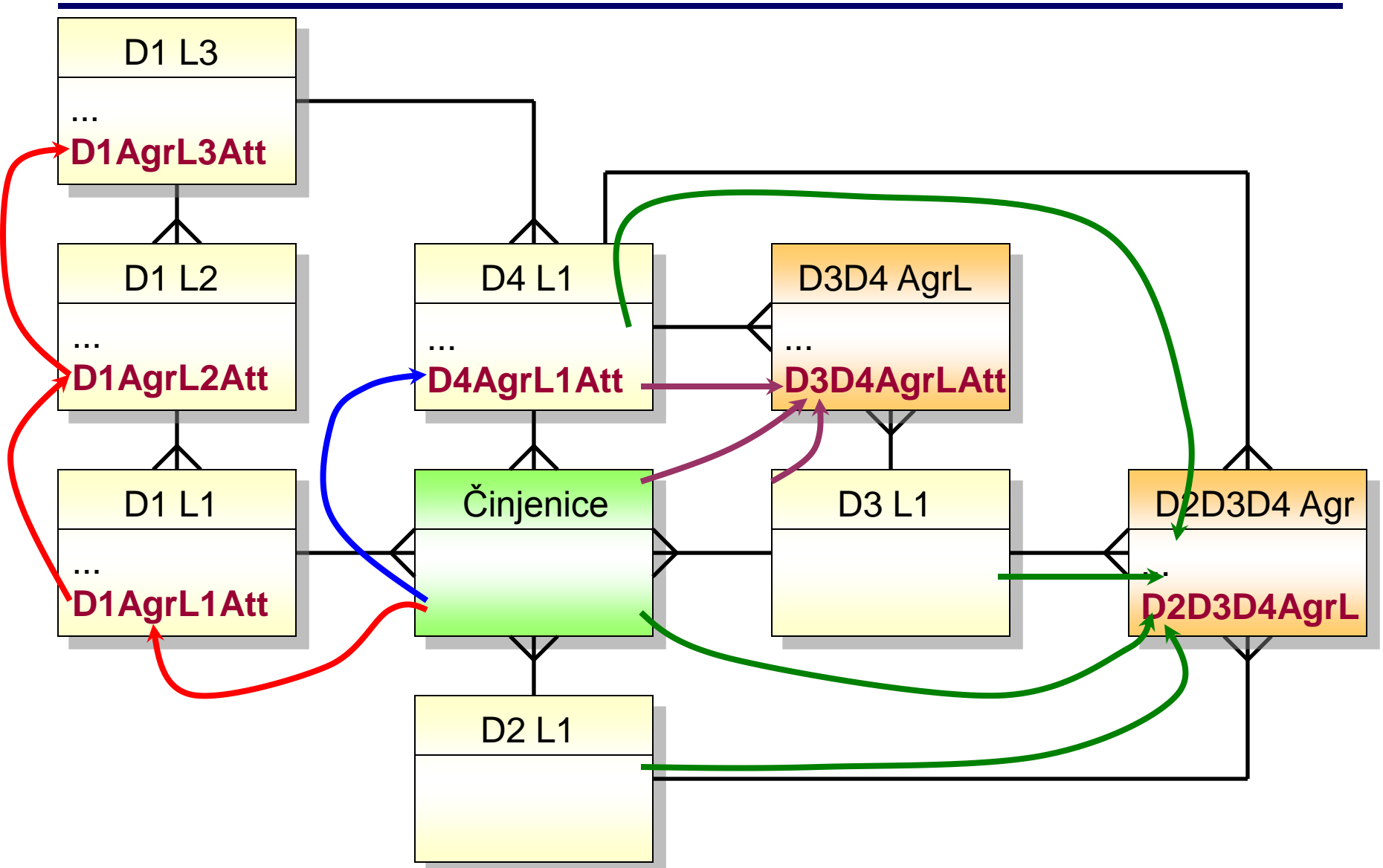
# Šema sa agregiranim podacima

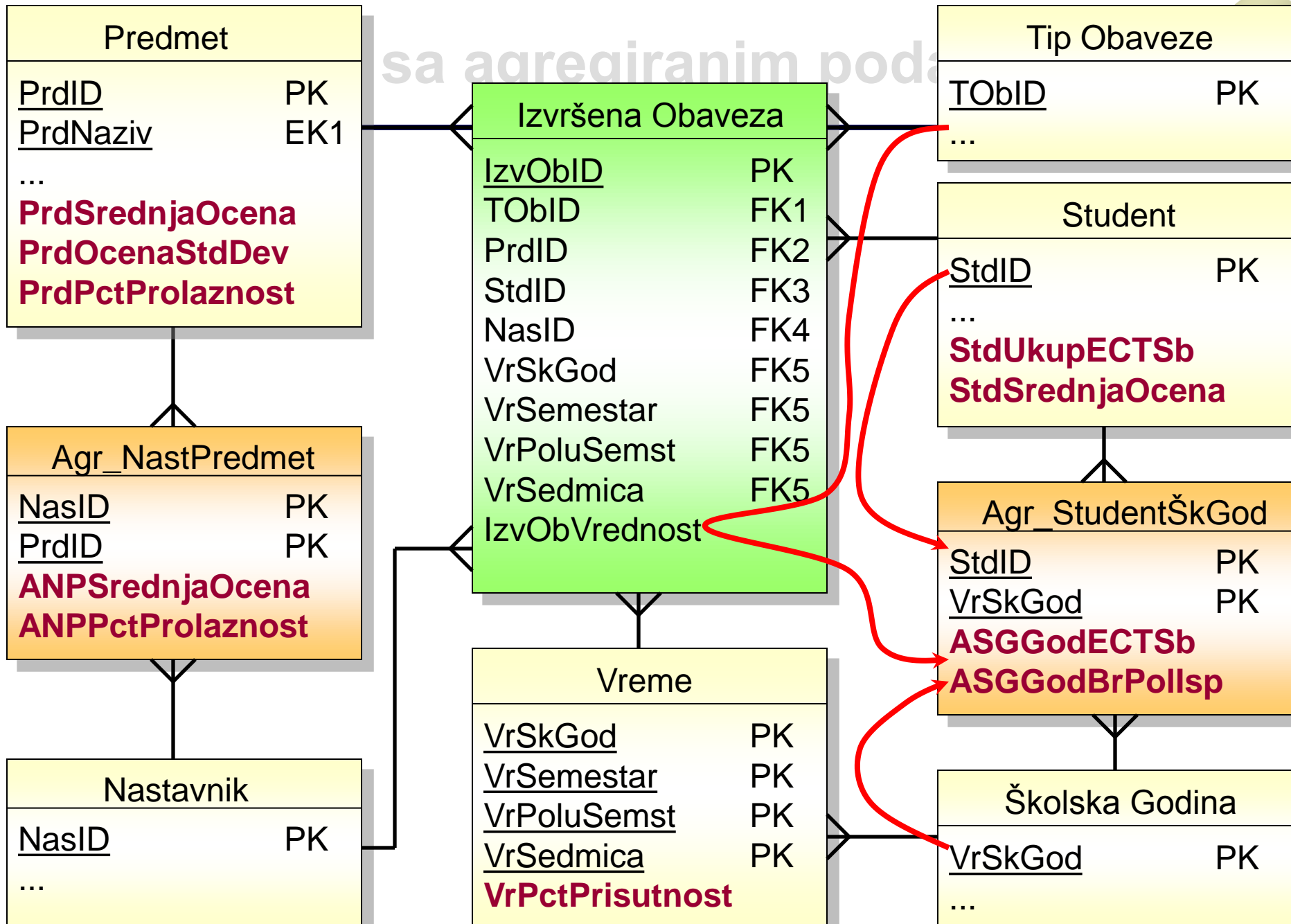
---

- **Agregirani (sumarni) podaci**
  - sumarni podaci po zadatim dimenzijama
    - redundantni, ali neophodni za efikasnu podršku različitih upita i analiza
    - čuvaju se u DW bazi podataka
    - generišu se (izračunavaju) u ECTL procesu
      - na osnovu formiranih činjeničnih podataka
      - često uz primenu skupovnih funkcija (SUM, COUNT, MIN, MAX, AVG, STDDEV,...)
- **Nivoi agregacije**
  - bilo koja kombinacija dimenzija
  - bilo koji nivo u hijerarhijskoj strukturi dimenzije



# Šema sa agregiranim podacima





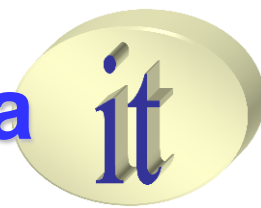
# Sadržaj

- Projektovanje DW sistema
- Projektovanje šeme BP
- Zvezdasta šema
- Strukture činjeničnih podataka
- Strukture dimenzionih podataka
- Hijerarhije dimenzija
- Šema tipa pahuljice
- Šema tipa sazvežđa
- Šema sa agregiranim podacima

# Pitanja i komentari

---





Kraj prezentacije

# Strukture šeme BP Data Warehouse sistema

---

*Logički i dimenzioni model  
DW sistema*