

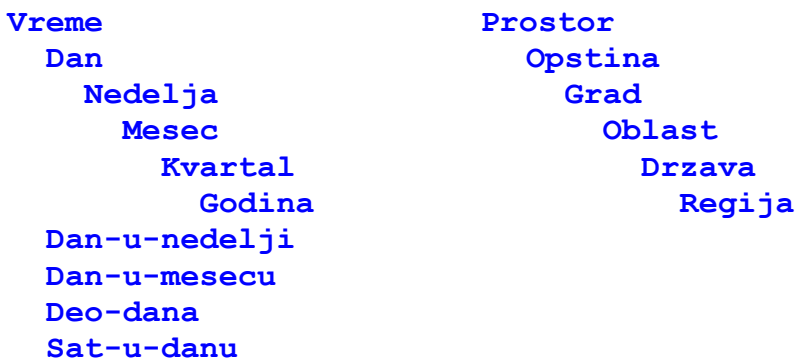
SKLADIŠTE PODATAKA – DIMENZIONI MODEL

Dimenziono modeliranje

Polazi se od toga da se atributi entiteta (tabela) mogu podeliti na dve grupe:

- ✓ **atributi mere**: atributi čija vrednost odražava meru (veličinu) nečega; bitna osobina je mogućnost svođenja;
- ✓ **atributi dimenzije**: atributi čija vrednost služi kao osnov klasifikacije i svođenja; mogu biti bez i sa varijabilnom granulacijom (nivoima svođenja).

Dva specijalna atributa dimenzije sa varijabilnom granulacijom: **Prostorni** i **Vremenski**



Prostorni dimenzioni atribut ima jednu šemu hijerarhije. Najniži nivo rezolucije je GPS geografska pozicija iz koje se dalje izvode viši nivoi rezolucije. U tom pogledu očekuje se uvođenje novog standardnog tipa podatka (POSITION?) i odgovarajućih standardnih funkcija izvođenja.

Vremenski dimenzioni atribut ima 5 paralelnih šema hijerarhije. Najniži nivo rezolucije je vreme (sat-minut-sekunda-...) i za to postoji standardni tip podatka TIMESTAMP. Postoji i čitav niz standardnih funkcija koje vrše izvođenje po raznim šemama hijerarhija (TIME, DATE, DAY_OF_WEEK, MONTH itd).

Kod dimenzionog modeliranja može istovremeno da postoji više vremenskih dimenzionih atributa (na primer, Dan-u-nedelji i Sat), pod uslovom da su one nezavisne, odnosno da nizu izvode jedna iz druge (kombinacija Dan i Dan-u-mesecu nema smisla).

Dve vrste tabela u dimenzionom modelu:

- ✓ **tabela fakata** (FT, jedna): sastoji se iz složenog primarnog ključa u koji ulaze ID-ovi svih dimenzija i jednog ili više atributa mere;
- ✓ **tabele dimenzija** (DTi, više): sastoje se iz prostog primarnog ključa koji odgovara jednoj komponenti primarnog ključa tabele fakata i jednog ili više atributa dimenzije.

Unutar jednog skladišta ili podskladišta podataka može se nalaziti više dimenzionalnih modela, od kojih neki mogu biti odvojeni a neki spregnuti preko zajedničkih dimenzija.

Dimenzije se u većini slučajeva javljaju kao par vrednosti:

- ✓ **ključ**: nema značenje i služi samo za pristup;
- ✓ **vrednost**: podatak koji ima značenje.

Izuzetak mogu biti vremenske dimenzije gde je isti podatak i ključ i vrednost.

Varijante dimenzionog modeliranja

Šema "zvezda" (star): jedan dimenzioni model

Sadrži jednu FT i za svaku dimenziju tačno jednu DT. DT mogu da sadrže denormalizovane podatke (neključne funkcijske zavisnosti koje nisu u 3. NF). Maksimalna dubina referisanja je 1.

$$FT \rightarrow DT_i$$

Šema "zvezda-pahuljica" (starflake): jedan dimenzioni model

Sadrži jednu FT i niz DT koje ne mogu da sadrže denormalizovane podatke, nego umesto toga sadrže reference na dodatne informacione tabele IT. Pri tome IT mogu da sadrže denormalizovane podatke. Maksimalna dubina referisanja je 2, pri čemu pojedine dimenzione tabele ne moraju imati reference na IT.

$$FT \rightarrow DT_i \rightarrow IT_i \\ DT_j$$

Šema "pahuljica" (snowflake): jedan dimenzioni model

Uz prethodni uslov ima i ograničenje da ni jedna IT ne sme da sadrži denormalizovane podatke, što dovodi do proizvoljne dubine referisanja

$$FT \rightarrow DT_i \rightarrow IT_{ia} \rightarrow IT_{ib} \dots \\ DT_j$$

Sve prethodno odnosilo se na pojedinačni dimenzioni model.

Šema sazvežđe (constellation): bar dva dimenziona modela.

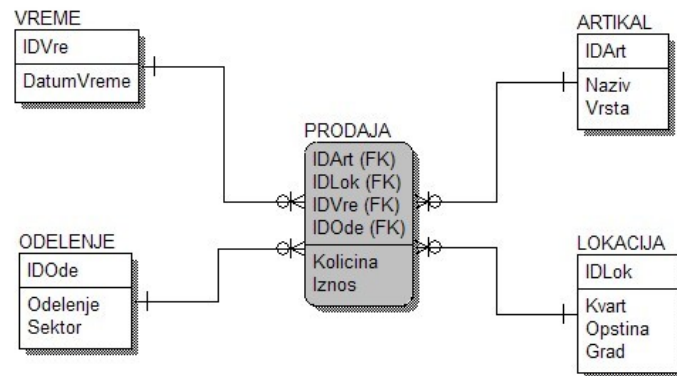
Situacija kada postoje bar dva dimenziona modela a njihove FT referišu bar jednu zajedničku DT. Pri tome dimenzioni modeli mogu biti bilo kog od prethodna tri tipa.

Prednost šeme "Zvezda" je smanjenje broja operacija spajanja, a mana povećanje memorijskog prostora.

Prednost ostalih šema je smanjenje memorijskog prostora, a mana povećanje broja operacija spajanja.

Primer: Prodaja artikala - šema "Zvezda"

Model podataka:



Skladište podataka:

PRODAJA (IDArt, IDLok, IDVre, IDOde, Kolicina, Iznos)

ARTIKAL (IDArt, Naziv, Vrsta)

Naziv->Vrsta

LOKACIJA (IDLok, Kwart, Opstina, Grad)

Kwart->Opstina, Opstina->Grad

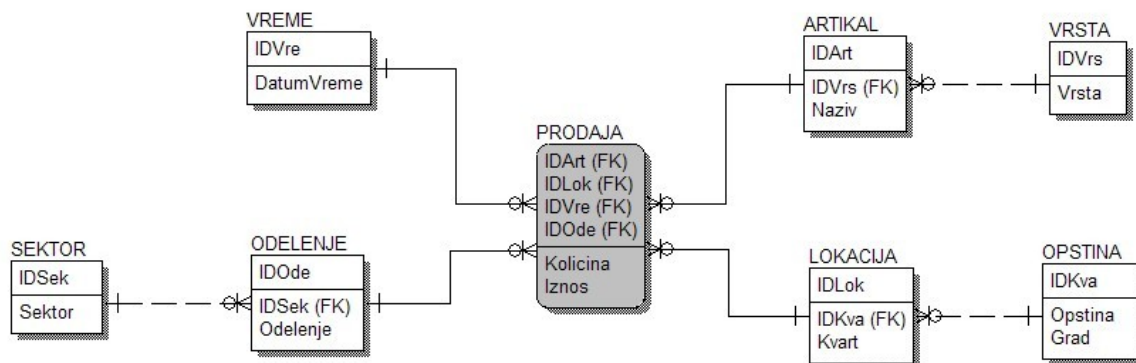
VREME (IDVre, DatumVreme)

ODELENJE (IDOde, Odeljenje, Sektor)

Odeljenje->Sektor

Primer: Prodaja artikala - šema "Zvezda-Pahuljica"

Model podataka:



Skladište podataka:

PRODAJA (IDArt, IDLok, IDVre, IDODE, Kolicina, Iznos)

ARTIKAL (IDArt, Naziv, IDVrs)

VRSTA (IDVrs, Vrsta)

LOKACIJA (IDLok, Kvart, IDOps)

OPSTINA (IDOps, Opstina, Grad)

Opstina->Grad

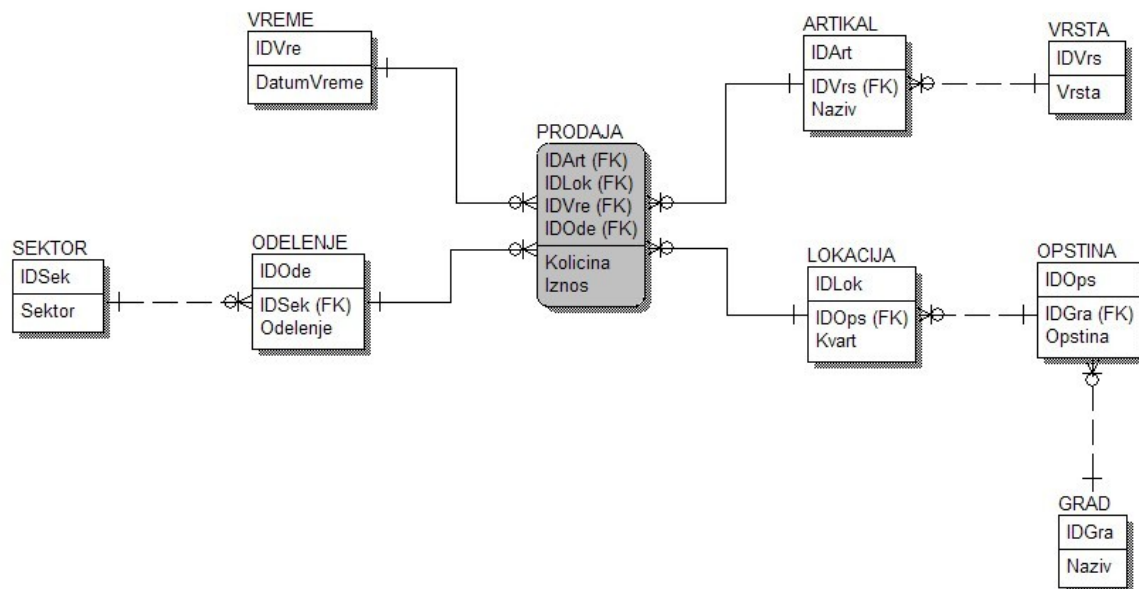
VREME (IDVre, DatumVreme)

ODELENJE (IDODE, Odeljenje, IDSek)

SEKTOR (IDSek, Sektor)

Primer: Prodaja artikala - šema "Pahuljica"

Model podataka



Skladište podataka:

PRODAJA (IDArt, IDLok, IDVre, IDOde, Kolicina, Iznos)

ARTIKAL (IDArt, Naziv, IDVrs)

VRSTA (IDVrs, Vrsta)

LOKACIJA (IDLok, Kvart, IDOps)

OPSTINA (IDOps, Opstina, IDGra)

GRAD (IDGra, Grad)

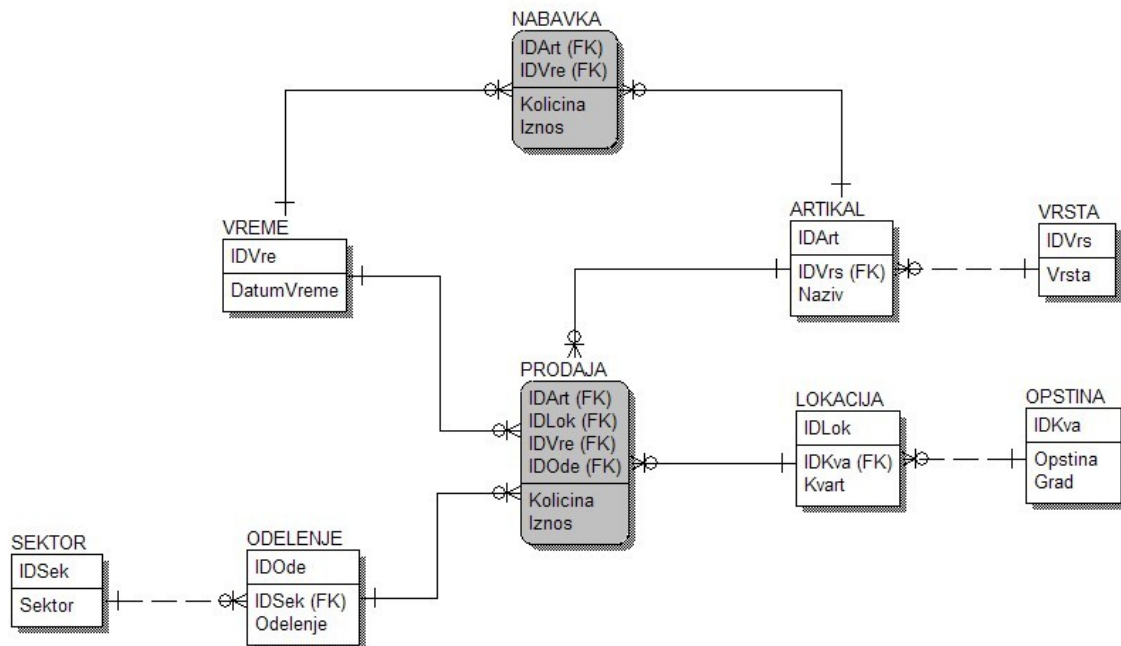
VREME (IDVre, DatumVreme)

ODELENJE (IDOde, Odeljenje, IDSek)

SEKTOR (IDSek, Sektor)

Primer: Prodaja i nabavka artikala - šema "Sazvežđe"

Model podataka:



Skladište podataka

PRODAJA (IDArt, IDLok, IDVre, IDOde, Kolicina, Iznos)

ARTIKAL (IDArt, Naziv, IDVrs)

VRSTA (IDVrs, Vrsta)

LOKACIJA (IDLok, Kvart, IDOps)

OPSTINA (IDOps, Opstina, Grad)

Opstina->Grad

VREME (IDVre, DatumVreme)

ODELENJE (IDOde, Odeljenje, IDSek)

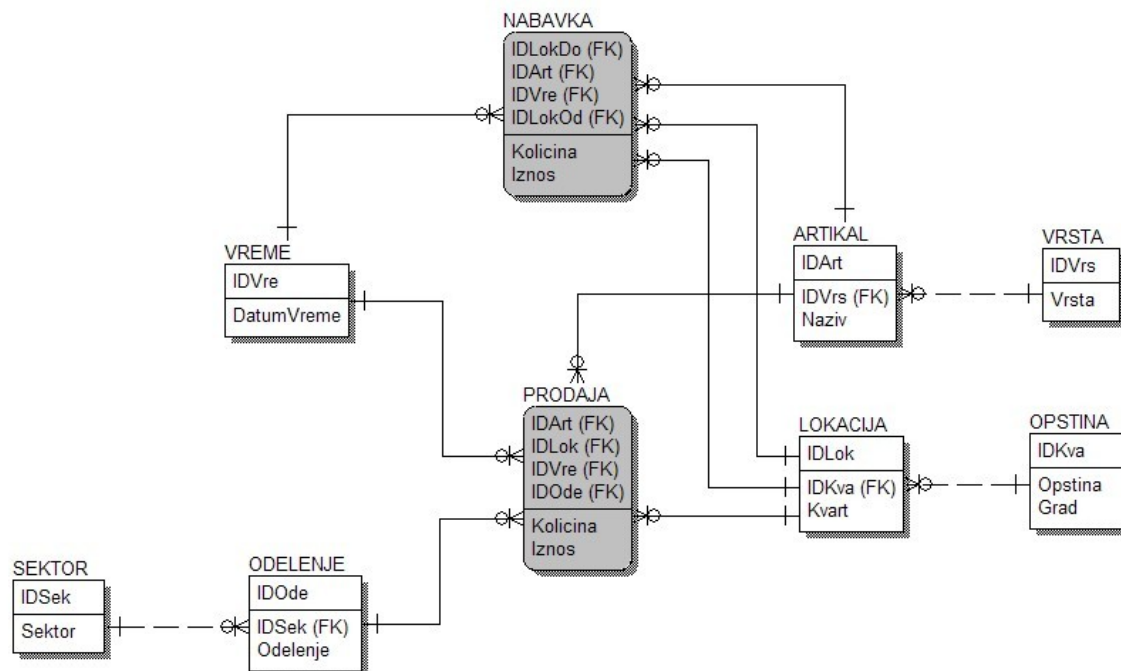
SEKTOR (IDSek, Sektor)

NABAVKA (IDArt, IDVre, Kolicina, Iznos)

Prodaja i nabavka su dve odvojene tabele fakata zato što evidentiraju nezavisne i odvojene poslovne događaje.

Primer: Prodaja i nabavka artikala - šema "Sazvežđe" - složenije

Model podataka:



Skladište podataka:

PRODAJA (IDArt, IDLok, IDVre, IDOde, **Kolicina**, **Iznos**)

ARTIKAL (IDArt, **Naziv**, IDVrs)

VRSTA (IDVrs, **Vrsta**)

LOKACIJA (IDLok, **Kvart**, IDOps)

OPSTINA (IDOps, **Opstina**, **Grad**)

Opstina->Grad

VREME (IDVre, **DatumVreme**)

ODELENJE (IDOde, **Odeljenje**, IDSek)

SEKTOR (IDSek, **Sektor**)

NABAVKA (IDArt, IDVre, IDLokOd, IDLokDo, **Kolicina**, **Iznos**)

Napomena:

U dimenzionom modelu otpreme dimenzija **lokacija** se javlja u dva svojstva (od i do) pa je referenca ka tabeli dimenzije **lokacija**.