

Napredne baze podataka - Prva domaća zadaća

Koristiti VCL sliku PMF_NBP_PSQL_olimpijada koja sadrži pripremljenu bazu podataka nbp_rdb u PostgreSQL bazi podataka; kako je opisano u pripremnom predavanju vezanom uz VCL. Shema nbp_rdb baze je na Merlinu.

Potrebno je izgraditi jednostavno skladište podataka temeljem podataka u nbp_rdb bazi podataka. Skladište treba biti tako modelirano da:

- pokriva vremenski period od najranijih do najkasnijih olimpijskih igara
- omogućava jednostavnu analizu nastupa i uspješnosti sportaša u vremenu prema disciplinama, sportovima, lokaciji igara i sportašima (spol, građa, godine, nacionalnost)

Zadatak je:

1. napisati što je činjenica koju skladište opisuje, uzevši u obzir i granulaciju podataka
2. definirati činjeničnu tablicu i pripadne dimenzijske tablice u zvjezdastom modelu skladišta (napisati CREATE TABLE naredbe za sve te tablice)
3. napraviti sliku zvjezdastog modela tog skladišta
4. napuniti dimenzijske tablice i činjeničnu tablicu podacima iz nbp_rdb baze podataka
5. izvršiti upite nad tablicama skladišta podataka kojima se mogu dobiti:
 - o top 10 najuspješnijih država na ljetnim i top 10 najuspješnijih država na zimskim OI
 - o top 10 najuspješnijih sportaša na ljetnim i top 10 najuspješnijih sportaša na zimskim OI nakon 1950. godine
 - o usporedbu broja disciplina u kojima su se žene natjecale kroz vrijeme
 - o prosječnu dob osvajača medalja po spolu u atletici i gimnastici (odvojeno) nakon 1970. godine
 - o istraživanje prednosti domaćeg terena – postotak medalja na pojedinim igrama koje su osvojili domaći sportaši u odnosu na ostale

Rješenje zadatka koje predajete na Merlin treba uključivati dvije datoteke:

- datoteku u tekstualnom (*plain text*) obliku koja sadrži:
 - o opis činjenice koju skladište opisuje (SQL komentar)
 - o SQL naredbe za kreiranje činjenične tablice i dimenzijskih tablica
 - o SQL naredbe za popunjavanje svih tih tablica temeljem podataka u nbp_rdb bazi
 - o SQL upite na tablice skladišta podataka koje daju odgovore na pitanja iz zadatka (točka 5)
- sliku zvjezdastog modela vašeg skladišta podataka (rukom crtanu ili generiranu iz alata npr. DataGrip)

Važne napomene:

- vremensku dimenziju se u ovom slučaju može pojednostavniti
- koristiti i izvedene mjere
- radi jednostavnosti pisanja upita, činjeničnu tablicu i dimenzijske tablice kreirati neposredno u nbp_rdb bazi
- sav SQL kod treba biti jasno komentiran!
- zadatak je moguće riješiti isključivo pisanjem SQL upita, nije potrebno koristiti nikakve druge programske jezike

- po potrebi koristiti postgresql tutorial (<https://www.postgresqltutorial.com/>) te pitati web (odlična rješenja za neke SQL upite ćete naći na stack overflow)