Univerzitet u Sarajevu

Elektrotehnički fakultet u Sarajevu

Stručni studij Razvoj softvera

Skladišta podataka, akademska 2024/2025

Predmetni profesor: R. prof. dr Samir Omanović dipl. ing. el.

Projekat skladišta podataka

Radili: Mahmutović Eldin 211-ST, Petković Marko 232-ST,

Pap Anand 238-ST, Skopljak Malik 181-ST

Uvod

Za izradu našeg projekta koristili smo:

1. SQLite DBMS (Database management system = sistem za upravljanje bazom podataka)
2. Već pripremljenu produkcijsku bazu podataka napunjenu početnim podacima, a sa kojom smo se bili upoznali kroz rad na predmetu UBP (Uvod u baze podataka, 1. semestar RS studija)
3. DB Browser aplikaciju za inspekciju stanja baze podataka i izvršavanje upita

Produkcijska baza se nalazi u trećoj normalnoj formi (3NF), te će biti potrebna denormalizacija prilikom prenosa podataka u DWH (Data warehouse = skladište podataka). U svrhu izvještavanja i analize, podaci se denormalizuju kako bi se unaprijedile performanse upita. Denormalizaciju ćemo obaviti spajanjima (join-ovima) odgovarajućih tabela.

Naš DWH smo osmislili tako da se sastoji od sljedećih dimenzionih modela zvjezdane strukture:

1. Model narudžbe i odgovarajuća mjera
2. Model inventure i odgovarajuća mjera

Za mjeru narudžbe vežemo četiri dimenzije:

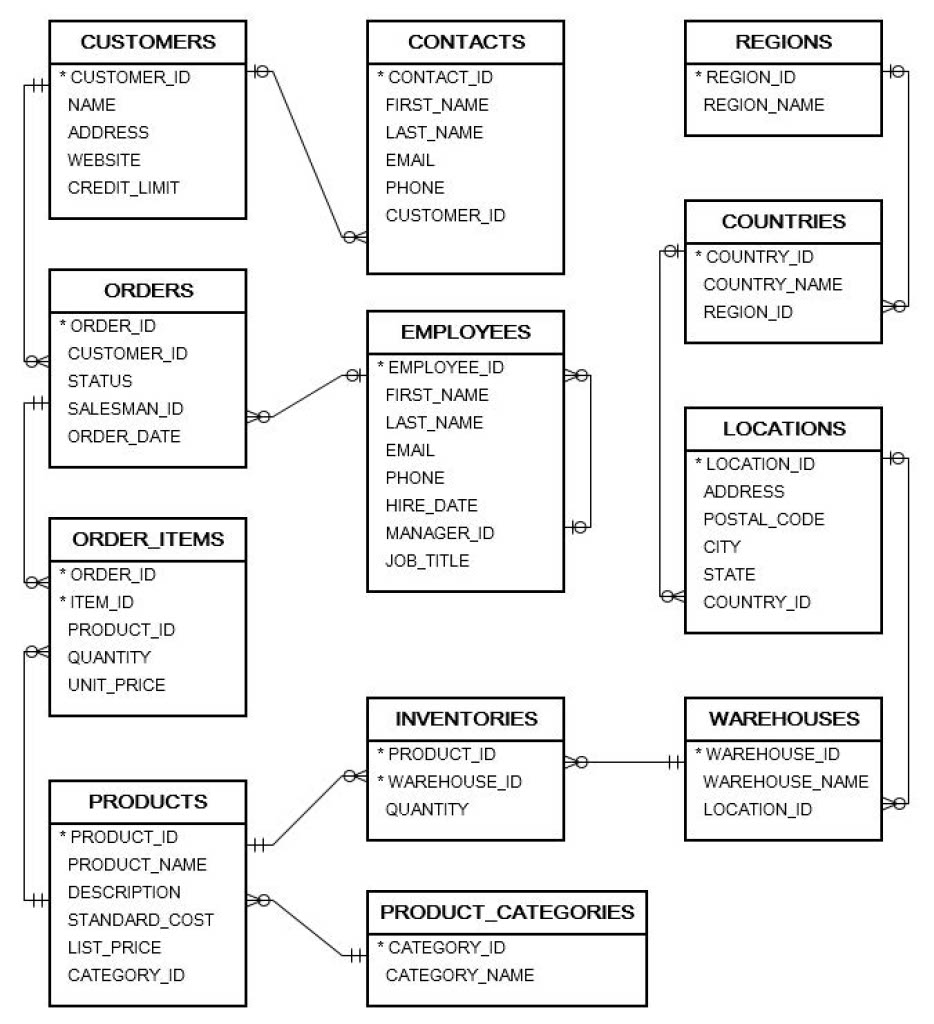
1. Uposlenik
2. Kupac
3. Proizvod
4. Vrijeme

A za mjeru inventura vežemo sljedeće dvije dimenzije:

1. Proizvod
2. Skladište

Dijagrami strukture produkcijske baze i DWH

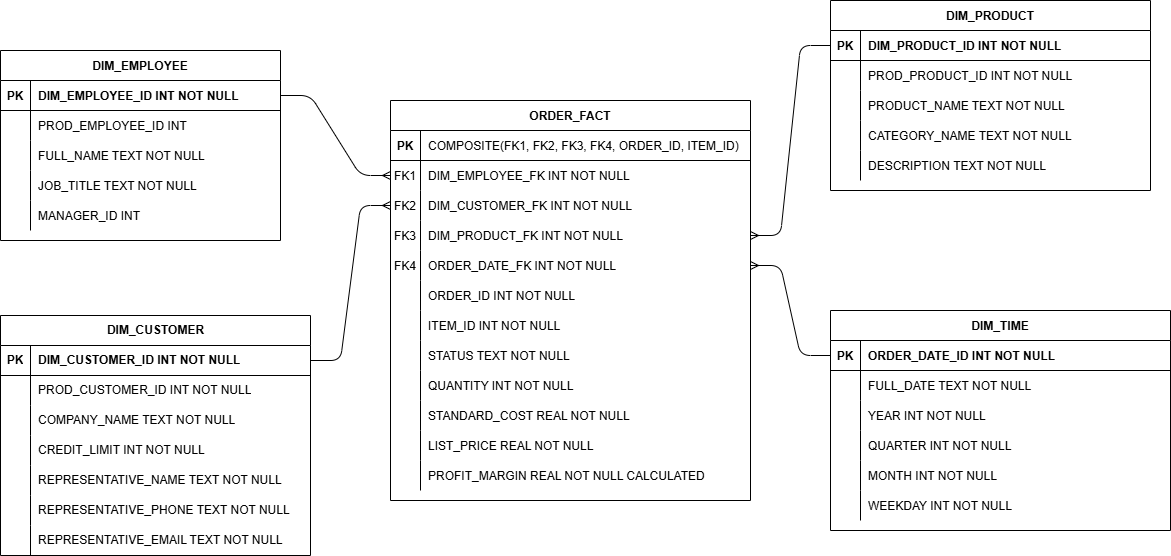
*ERD (Entity-relationship dijagram) produkcijske baze*



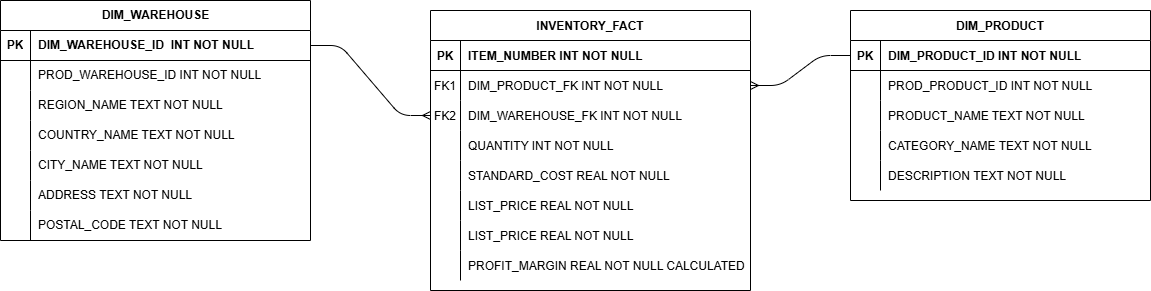
*Matrica arhitekture DWH*



*ERD ORDER\_FACT tabele*



*ERD INVENTORY\_FACT tabele*



Kreiranje DWH

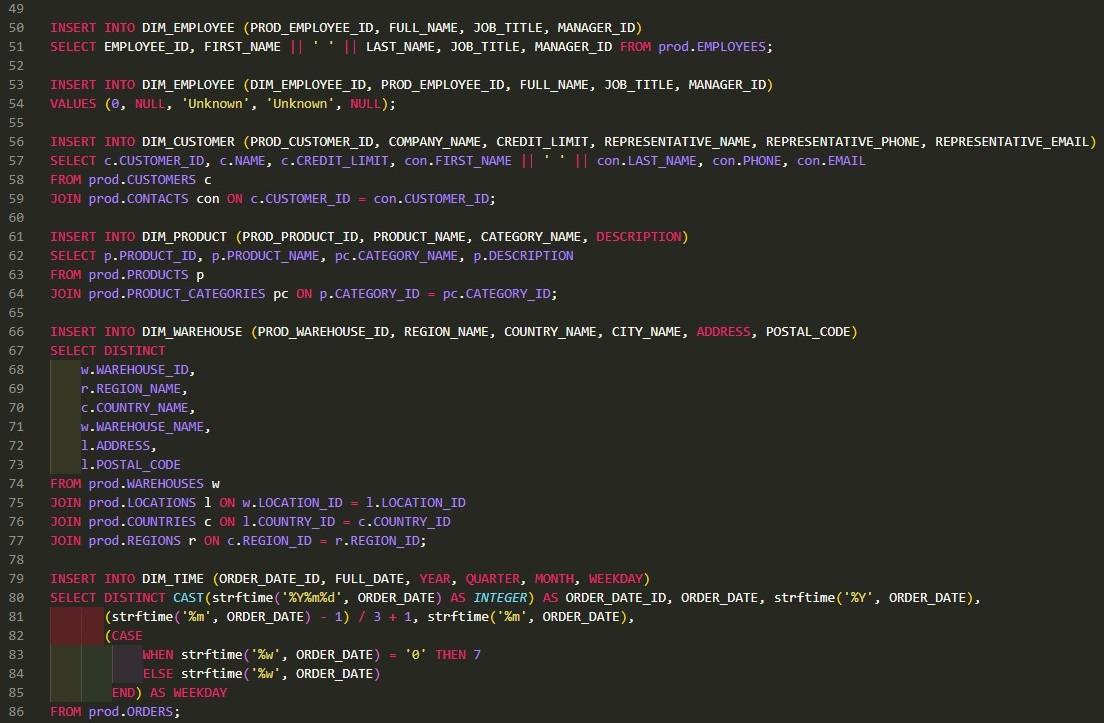
Na početku ćemo spojiti produkcijsku bazu na naš novoformirani DWH, te postaviti obavezu poštovanja ograničenja stranih ključeva.



Narednim komandama kreiramo redom dimenzije koje ćemo kasnije vezati na naše fact tabele.



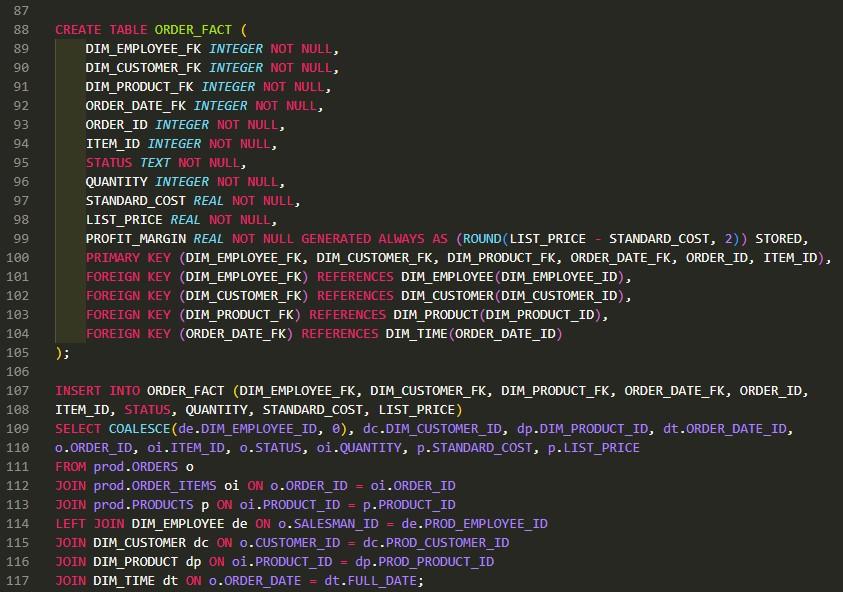
*Punjenje tabela dimenzija*



Obratite pažnju da smo u **DIM\_EMPLOYEE** dimenziju unijeli zaseban red koji označava uposlenika koji nije zabilježen tokom dogovaranja narudžbe (njegov ID je vrijednosti NULL). Da nismo tako postupili, prilikom unosa podataka u **ORDER\_FACT** tabelu, narudžbe koje su dogovorene od strane prodavača koji nisu zabilježeni bi bile jednostavno odstranjene i ne bi bile uzete u razmatranje za analize koje slijede nakon kreiranja DWH, a time ne bi dobili pravu sliku stanja.

Dodatna napomena: kupac je kompanija, a ne pojedinac i sa njom naša firma komunicira preko kontakt osobe koja je predstavnik te kompanije sa kojom poslujemo.

*Kreiranje i punjenje ORDER\_FACT tabele mjere*



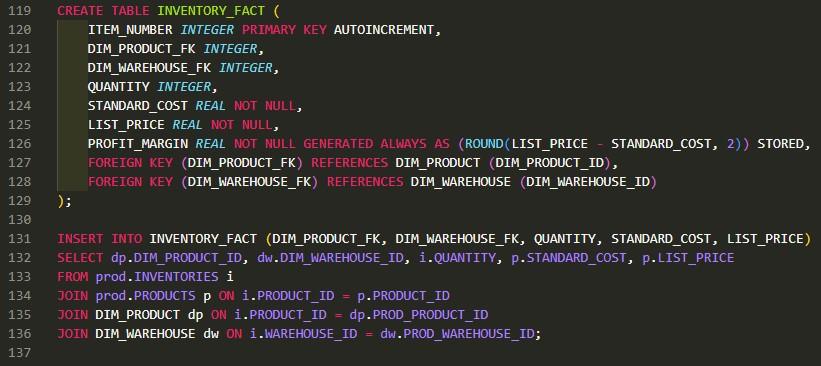
Činjenice naše **ORDER\_FACT** tabele su: **QUANTITY** (količina naručenog proizvoda), **STANDARD\_COST** (trošak proizvodnje datog proizvoda iz stavke), **LIST\_COST** (cijena koju potražujemo prilikom narudžbe) i **PROFIT\_MARGIN** (marža ili očekivana zarada) izračunata na osnovu razlike cijene i troška proizvodnje.

Ovdje je bitno obratiti pažnju da **ORDER\_ID** i **STATUS** kolone predstavljaju degenerisane dimenzije. To su podaci koji liče na dimenzije samo što se ne referenciraju preko stranog ključa (FK) već su direktno upisani u tabelu mjere. Nisu prave dimenzije, a po svome smislu, one to jesu.

**ORDER\_ID** je redni broj narudžbe, a **STATUS** uzima vrijednosti “Canceled”, “Pending” ili “Shipped”, ovisno o stanju date narudžbe.

Problem narudžbe čiji prodavač nije poznat je riješen referenciranjem na specijalni red iz **DIM\_EMPLOYEE** dimenzije čiji je **DIM\_EMPLOYEE\_ID** jednak **0**. Time je sačuvan zapis o narudžbi, a uposlenik koji je nju dogovorio je označen kao “Unknown” (nepoznat).

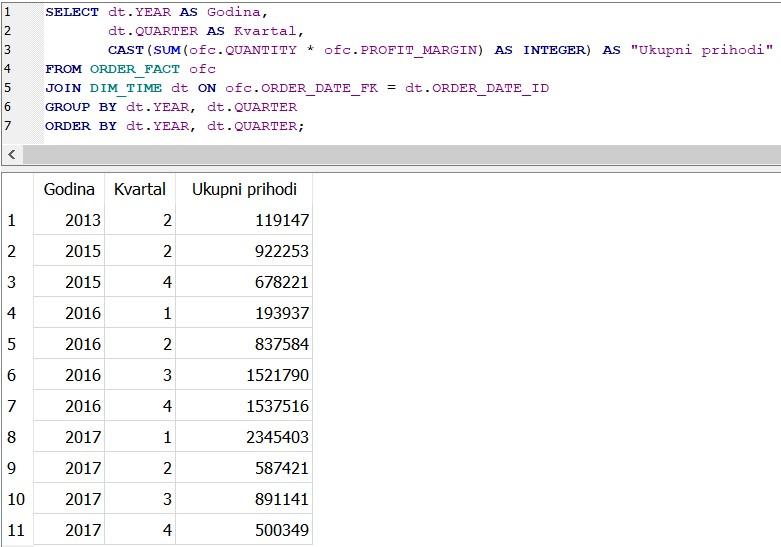
*Kreiranje i punjenje INVENTORY\_FACT tabele mjere*



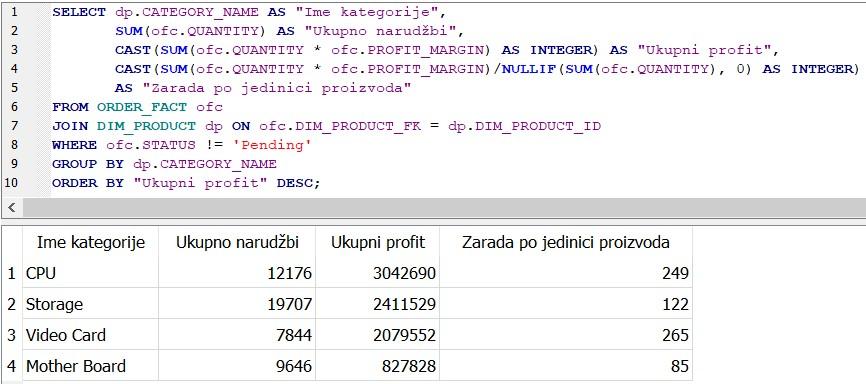
Činjenice **INVENTORY\_FACT** tabele su: **QUANTITY** (količina proizvoda na stanju u datom skladištu), **STANDARD\_COST** (trošak proizvodnje), **LIST\_PRICE** (prodajna cijena) i **PROFIT\_MARGIN** (marža), slično kao za **ORDER\_FACT** tabelu mjere.

Primjeri upita nad ORDER\_FACT tabelom iz DWH

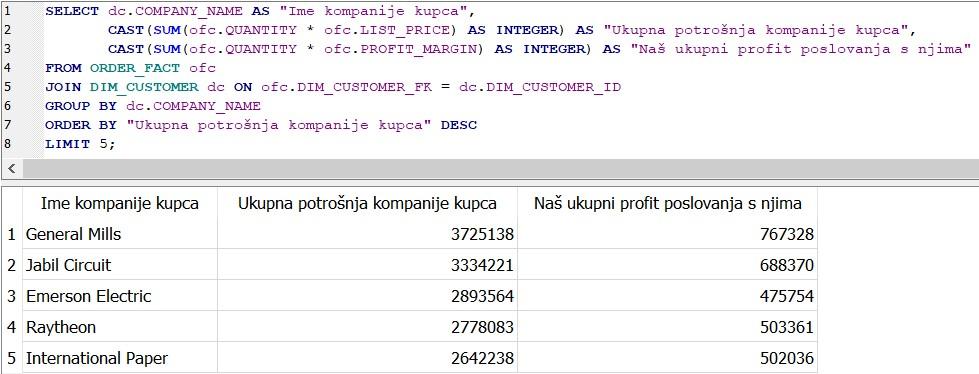
1. Prihodi po kvartalima



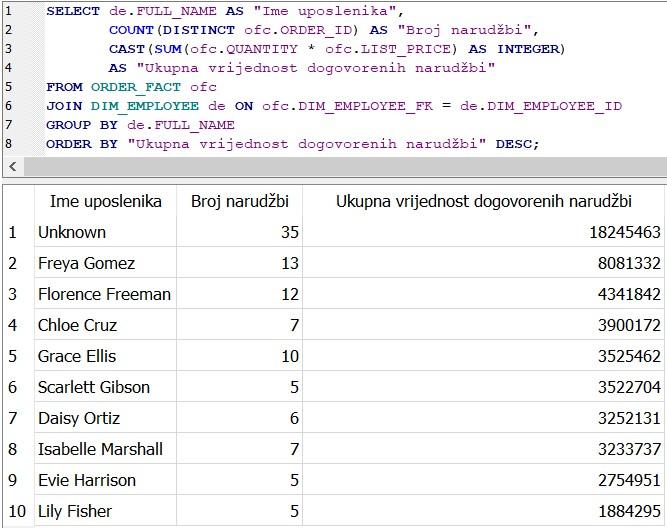
2. Ostvareni profit po kategoriji proizvoda



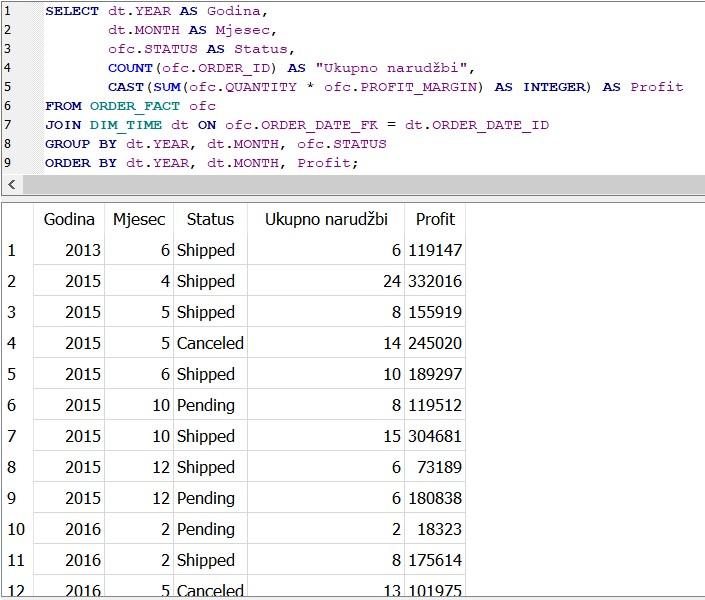
3. Top 5 kompanija kupaca po ukupnoj potrošnji



4. Uspjeh uposlenika prodavača rangirano opadajuće

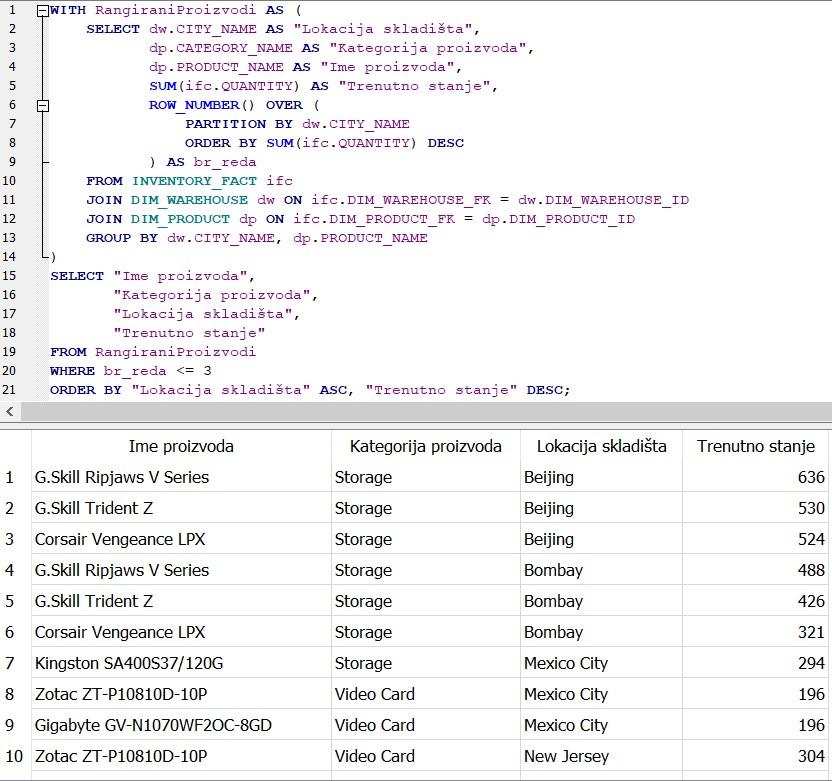


5. Mjesečni trendovi po statusu narudžbi

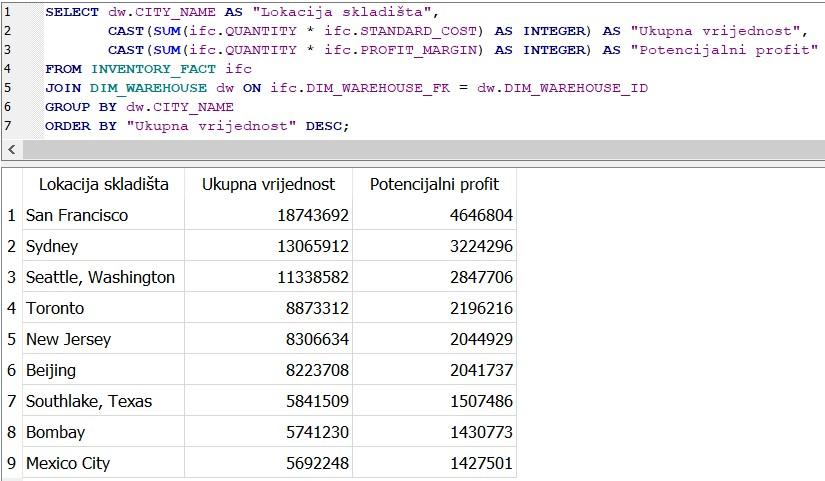


Primjeri upita nad INVENTORY\_FACT tabelom iz DWH

1. Top 3 proizvoda na stanju po različitim skladištima



2. Skladišta sa najvećom vrijednošću



3. Top 10 najzastupljenijih proizvoda u skladištima

