****

Fakultet tehničkih nauka

Inženjerstvo informacionih sistema

Predmet: Projektovanje skladišta podataka

Projekat

***Regionalna prodaja u SAD***

Student

**Marina Cvijetić IT60-2019**

Sadržaj

[Zadatak i ciljevi projekta 3](#_Toc125488281)

[Opis postupka projektovanja DW sistema 3](#_Toc125488282)

[Specifikacija zahteva korisnika 4](#_Toc125488283)

[Specifikacija modela 5](#_Toc125488284)

[1. Specifikacija izvora podataka 6](#_Toc125488285)

[2. Specifikacija ciljanog DataWarehouse sistema 7](#_Toc125488286)

[2.1 Specifikacija zahtevanih dimenzija 7](#_Toc125488287)

[2.2 Specifikacija zahtevanih mera 8](#_Toc125488288)

[Opis ETL procesa 9](#_Toc125488289)

[1. Kreiranje dimenzija i tabele činjenica 9](#_Toc125488290)

[2. Punjenje tabela dimenzija 15](#_Toc125488291)

[3. Punjenje tabele vremenske dimenzije 23](#_Toc125488292)

[4. Punjenje tabele činjenica 28](#_Toc125488293)

[Izveštaji 36](#_Toc125488294)

[Prvi izveštaj 36](#_Toc125488295)

[Drugi izveštaj 38](#_Toc125488296)

[Treći izveštaj 39](#_Toc125488297)

[Četvrti izveštaj 41](#_Toc125488298)

[Peti izveštaj 42](#_Toc125488299)

[Zaključak 44](#_Toc125488300)

# **Zadatak i ciljevi projekta**

Skladištenje podataka je usmereno na pravovremeno dobijanje relevantnih informacija potrebnih u postupcima odlučivanja. Skladište podataka je moćan sistem za podršku odlučivanju, s toga je zadatak ovog projekta da se omogući pregled poslovanja, odnosno uspešnost prodaje trgovinskih radnji u Sjedinjenim Državama na regionalnom nivou. Cilj je da se procesom integracije podataka kreira skladište podataka iz kojeg krajnji korisnici mogu sprovoditi ad-hock analize podataka i praviti izveštaje vezane za prodaje u različitim regijama Sjedinjenih Država.

# **Opis postupka projektovanja DW sistema**

Analiza i specifikacija korisničkih zahteva je prvi korak prilikom projektovanja skladišta podataka. Ključan korak jer zahteva vrlo detaljno razumevanje poslovanja i tok informacija u okviru njega, kako bi se identifikovali ciljevi, kritični faktori uspeha, ali i ustanovili procesi poslovanja, nadležnosti, poslovna pravila.

Projektovanje šeme DW baze podataka je naredni korak i on obuhvata:

1. Projektovanje konceptualne šeme BP
2. Projektovanje implementacione šeme BP
3. Projektovanje fizičke organizacije šeme BP

Nakon projektovanja šeme DW baze podataka sledi projektovanje arhitekture DW sistema u odnosu na prethodno izmodelovane dijagrame i logičke modele, odnosno konačna realizacija fizičke strukture skladišta podataka.

Projektovanje ETL softverske podrške podrazumeva odabir alata, odnosno softvera u okviru kog će se vršiti ETL proces. On podrazumeva ekstrakciju podataka iz izvornih sistema, njihovu transformaciju u oblik pogodan za analizu i donošenje odluka, a zatim i njihovo učitavanje u ciljano skladište podataka.

Poslednji korak jeste projektovanje softverske podrške za izveštavanje i analizu podataka. DW sistem ima za cilj da obezbedi tabelarni i grafički prikaz podataka u skladu sa upitom, odnsono zahtevom korisnika čija je svrha da pruže neophodne informacije kao podrške u daljim procesima odlučivanja.

# **Specifikacija zahteva korisnika**

Analizom specifikacije zahteva korisnika, utvrđeno je na koja pitanja korisnik želi da dobije odgovor, odnosno statistike i poslovni rezultati u odnosu na određene faktore i u određenim vremenskim periodima. Odgovori na pitanja i rezultati statistike dobijeni su detaljnom analizom podataka i oni su sledeći:

1. Preko kog kanala prodaje je ostvaren najveći prihod u Kaliforniji, Tenesiju, Koloradu i Floridi u 2019. godini?
   * Kalifornija – In-Store
   * Tenesi – Online
   * Kolorado – In-Store
   * Florida – In-Store
2. Grafički prikaz linijskog kretanja prihoda prodavnica po regijama u periodu od marta do jula 2020. godine?
   * Na srednjem zapadu od marta do početka aprila zabeležen je blagi porast prihoda sa 7 na blizu 9 miliona dolara. Od sredine aprila do početka juna prihod kontinualno opada i u proseku iznosi nešto više od 6 miliona, a zatim se od juna do jula dešava nagli porast do čak 11 miliona dolara.
   * Na severoistoku je zabeležen pad prihoda od marta do sredine aprila sa skoro 9 miliona na blizu 6 miliona dolara. Stanje ostaje isto sve do kraja maja i početka juna kada se dešava mali porast na nešto malo više od 6 miliona. Ipak sredinom juna, prihod opet opada ispod 6 miliona dolara.
   * Na jugu je zabeležen nagli rast prihoda u periodu od marta do početka aprila sa 5 na nešto više od 8 miliona. Tokom aprila i maja, kontinualno opada ispod 6 miliona, a zatim se početkom juna dešava blagi rast.
   * Na zapadu je zabeležen nagli rast prihoda od marta do početka maja sa nešto više od 5 miliona na više od 10, a zatim se od maja do jula dešava pad na malo više od 6 miliona.
3. Koliko proizvoda na popustu većem od 10% je prodao svaki tim prodaje u mesecu septembru, oktobru, novembru i decembru za svaku državu?
   * Svrha ovog izveštaja jeste da se prati uspešnost rada timova prodaje. Odgovor na ovo pitanje prikazan je u poglavlju *Izveštaji* (*Slika. 12)* zbog prevelike količine informacija.
4. U kojim državama je ostvaren najveći broj porudžbina? (*Data set* kreiran za potrebe ovog upita omogućava različite varijacije upita. Može da pruži odgovor i na pitanja koliki je srednji prihod po državama i kolika je gustina naseljenosti u svakoj državi.)
   * Najviše porudžbina je realizovani na teritoriji Teksasa i Kalifornije.
5. Top pet proizvoda najprodavanijih proizvoda za svaku godinu?
   * 2018. godina – Ornamenti, aksesoari, tanjiri, trpezarijski nameštaj, čaše za koktele i skulpture
   * 2019. godina – Ormani, tanjiri, telefoni, ćebad, aksesoari
   * 2020. godina – Poslužavnici, venci, suveniri, komode, zidni okviri

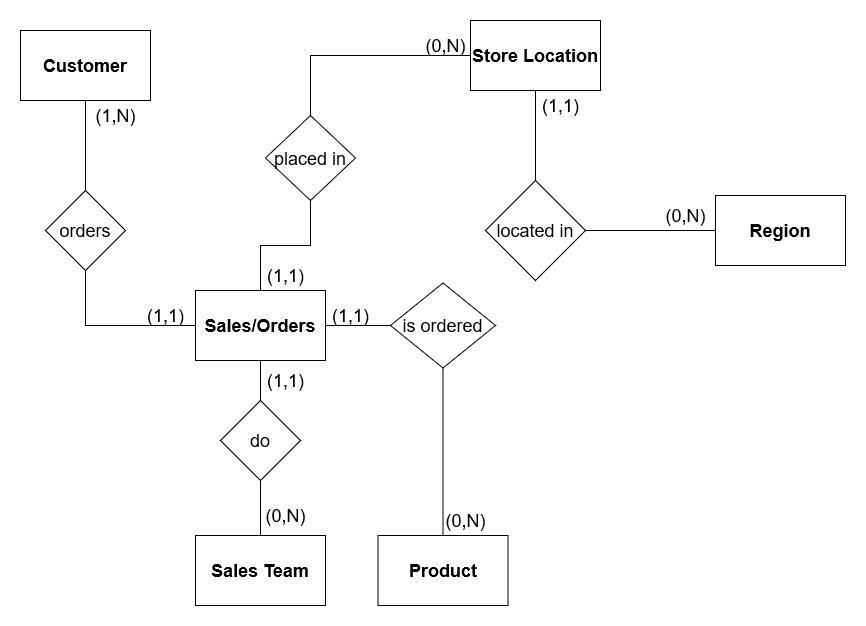
# **Specifikacija modela**

## **Specifikacija izvora podataka**

Set podataka za potrebe projekta je preuzet sa sajta DataWorld i podaci se odnose na regionalnu prodaju na teritoriji Sjedinjenih Država u periodu od 2018. godine do 2020. godine.

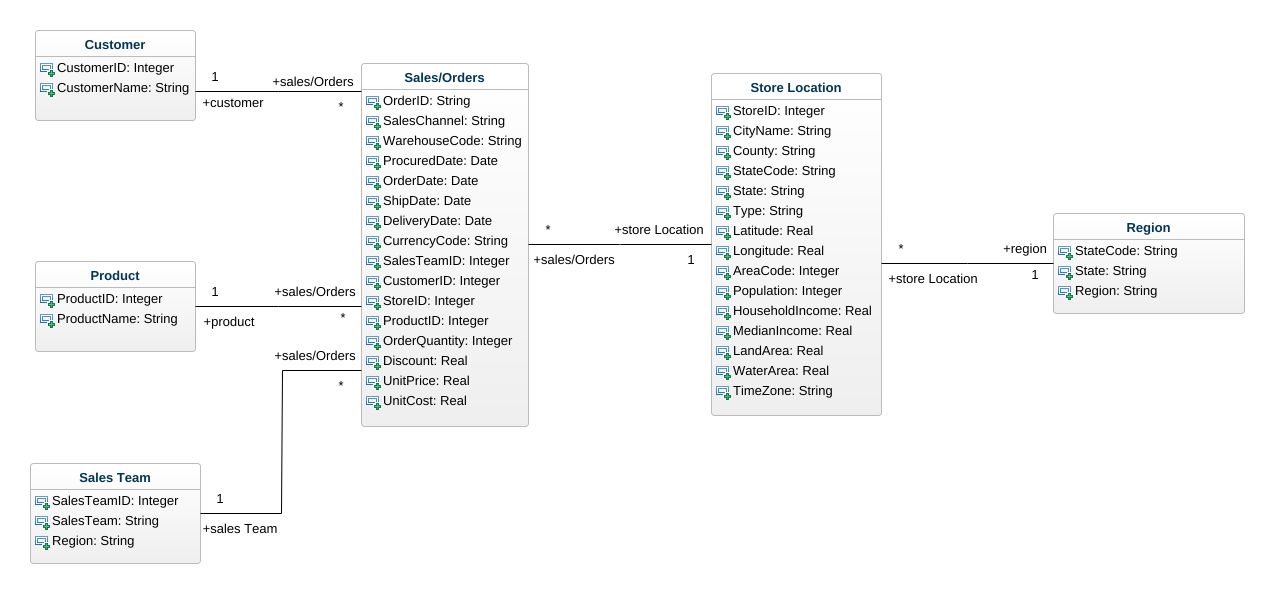
Set podataka predstavlja relacionu bazu podataka (OLTP) sačinjenu od 6 relacija, odnosno šest tabela koje sadrže podatke o lokaciji prodavnice (*StoreLocation*), timu prodaje (*SalesTeam*), izvršenim porudzbinama i prodajama (*Sales*), državi (*Region*), proizvodu (*Product*) i kupcima (*Customer*).

Radi boljeg razumevanja procesa i toka podataka, na sledećoj slici biće prikazan EER model izvorne transakcione baze podataka. (*Slika 1*)



Slika 1. EER Dijagram izvorne relacione baze podataka US Regional Sales.

Takođe je prikazan i dijgram klasa kako bi se pružio detaljniji pregled strukture šema relacija u izvornoj bazi podataka. (*Slika 2*)



Slika 2. Class Dijagram izvorne relacione baze podataka US Regional Sales

## **Specifikacija ciljanog DataWarehouse sistema**

### **2.1 Specifikacija zahtevanih dimenzija**

Na osnovu izvorne OLTP šeme kreirana je OLAP šema DW baze podataka. Za izradu OLAP šeme, podatke je najpre potrebno rasporediti u hijerarhijske grupe, odnosno definisati dimenzije. Osim dimenzija, važno je definisati proces koji će se posmatrati kroz tabelu činjenica u skladu sa specifikacijom zahteva korisnika.

Dimenzije koje su definisane su:

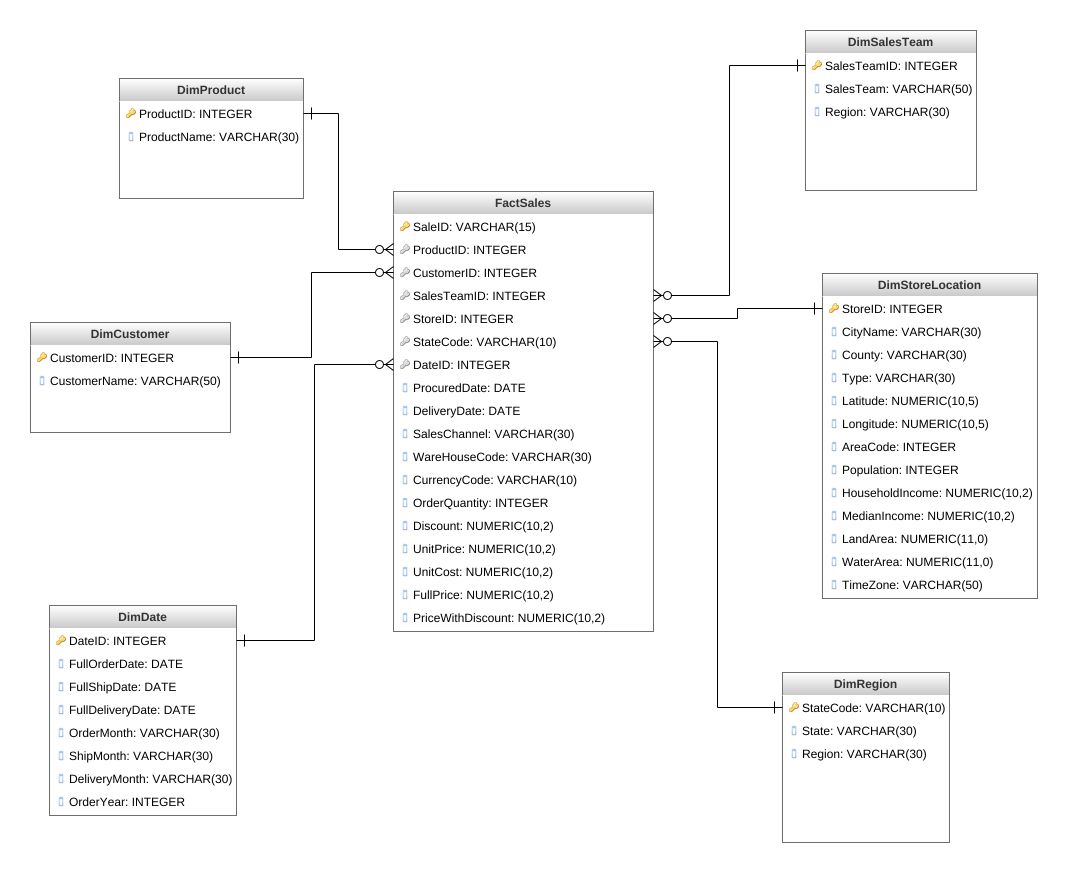
* dimCustomer
* dimProduct
* dimRegion
* dimStoreLocation
* dimSalesTeam
* dimDate

Proces koji će se posmatrati je proces prodaje, iz tog razloga tabela činjenica je kreirana od tabele *Sales*.

### **2.2 Specifikacija zahtevanih mera**

U tabeli činjenica implementirane su dve dodatne mere, *FullPrice* i *PriceWithDiscount*. *FullPrice* se dobija množenjem količine kupljenog proizvoda sa cenom po jedinici proizvoda kako bi se dobila ukupna cena po porudzbini. *PriceWithDiscount* se računa prvo računanjem iznosa popusta, a zatim odbijanjem istog od ukupne cene. Struktura skladišta je promenljiva u odnosu na zahteve, te je moguće vršiti određene izmene u zavisnosti od budućih potreba korisnika.

Prilikom modelovanja OLAP šeme, cilj je bio da se postigne denormalizovana struktura, odnosno šema zvezde. Postojala je i mogućnost da se realizuje šema pahulje tako što bi dimenzija *dimRegion* bila povezana sa *dimStoreLocation*, kao i u izvornoj OLTP šemi. Ipak, podaci o državama su odvojeni u posebnu dimenziju koja je povezana direktno sa tabelom činjenica zbog potrebe grupisanja podataka u odnosu na države. U nastavku je prikazana OLAP zvezda šema *US Regional Sales* skladišta podataka. (*Slika 3*)



Slika 3. OLAP zvezda šema US Regional Sales skladišta podataka

# **Opis ETL procesa**

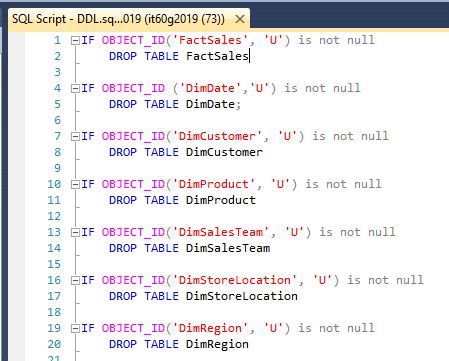
## **Kreiranje dimenzija i tabele činjenica**

Nakon što je definisana logika i struktura, izmodelovan dimenzioni model skladišta podataka, potrebno je kreirati tabele dimenzija i činjeničnu tabelu u koje će biti smešteni prethodno ekstraktovani i transformisani podaci iz izvorne relacione baze podataka.

Tabele dimenzija i tabela činjenica kreirane su putem SQL DDL naredbi. Kreiran je konačan SQL skript koji je prethodno pokrenut u okviru alata *Microsoft SQL Server Management Studio* kako bi se utvrdilo da kod ne sadrži greške.

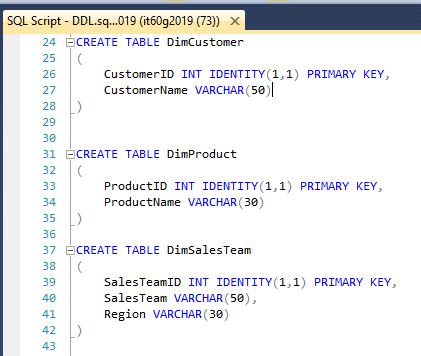
SQL skript za kreiranje biće prikazan u narednih nekoliko slika.

Na početku skripta je potrebno izvršiti proveru da li već postoje kreirani objekti u bazi pod istim nazivom, i ukoliko postoje – obrisati ih. (*Slika 4*)

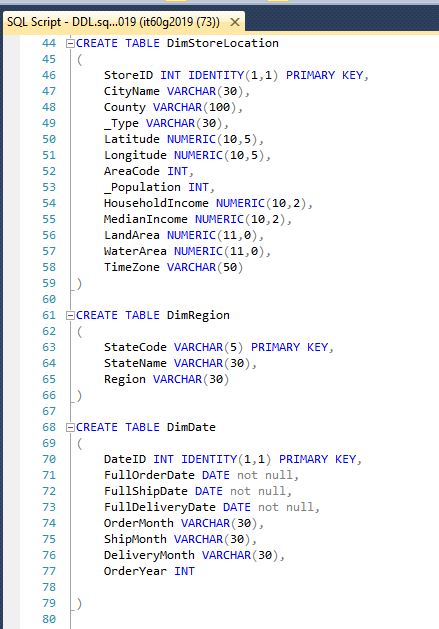


Slika 4. SQL Skript – Provera/Brisanje objekata

U nastavku sledi prikaz DDL naredbi za kreiranje tabela dimenzija. (*Slika 5-6*)

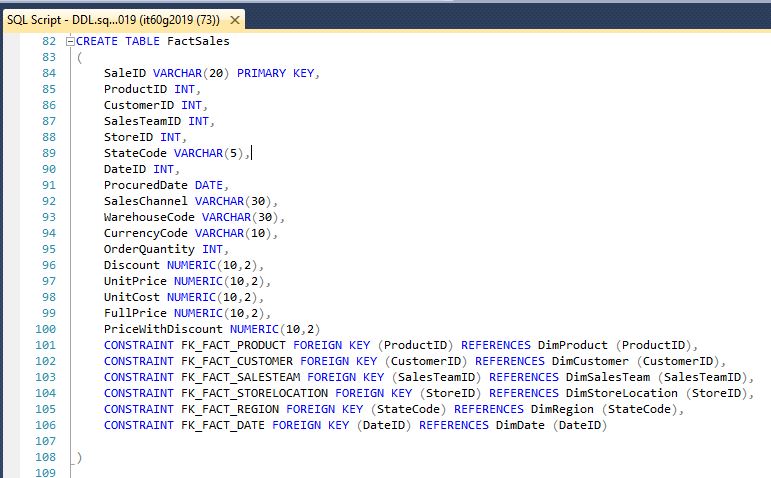


Slika 5. SQL Skript – Kreiranje tabela dimenzija 1



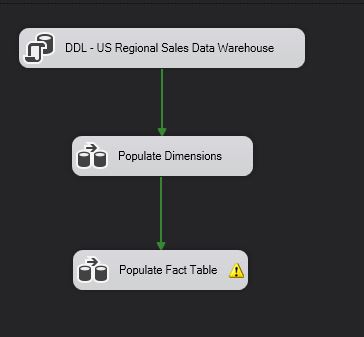
Slika 6. SQL Skript – Kreiranje tabela dimenzija 2

Tabela činjenica se kreira u istom *Data Flow-*u, ali poslednja zbog ograničenja referencijalnog integriteta, odnosno stranih ključeva koje sadrži i zato se nalazi na kraju SQL skripta, kao poslednja naredba koja će se izvršiti. (*Slika 7*)



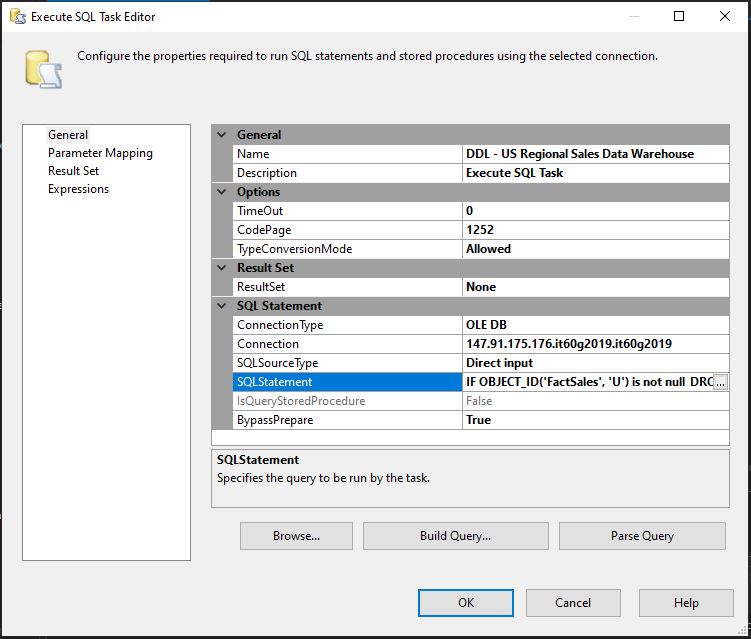
Slika 7. SQL Skript – Kreiranje tabele činjenica

SQL skript za kreiranje se izvršava u okviru ETL procesa, kao prvi korak u *Control Flow-*u realizovan putem komponente *Execute SQL Task.* Komponenta *Execute SQL Task* je preimenovana u „DDL – US Regional Sales Data Warehouse“*.* (*Slika 8*)

**

Slika 8. Prikaz ETL procesa u kartici Control Flow

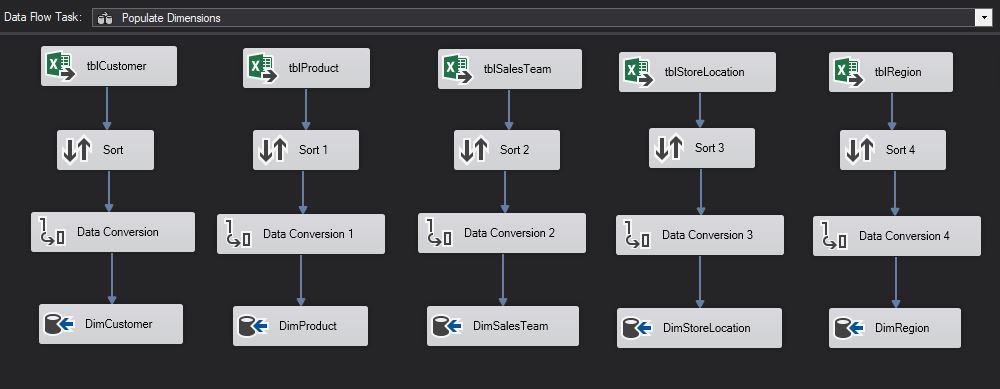
Komponenta *Execute SQL Task* zahteva dodatna podešavanja. Potrebno je izvršiti konekciju na server na kom će se kreirati DW baza podataka unosom konekcionog stringa u polja *Connection* i kopirati prethodno pripremljeni SQL skript u selektovano polje *SQLStatement*. (*Slika 9*)



Slika 9. Podešavanja u komponenti Execute SQL Task

## **Punjenje tabela dimenzija**

Kao što je ranije prikazano u kartici *Control Flow*, ETL proces se nakon izvršavanja SQL skripta nastavlja iniciranjem sledeće komponente, a to je *DataFlow* koja je preimenovana u “Populate Dimensions” (*Slika 8*). Ulaskom u navedeni *DataFlow* započinje se proces popunjavanja prethodno kreiranih tabela dimenzija, odnosno proces punjenja skladišta podataka. (*Slika 10*)



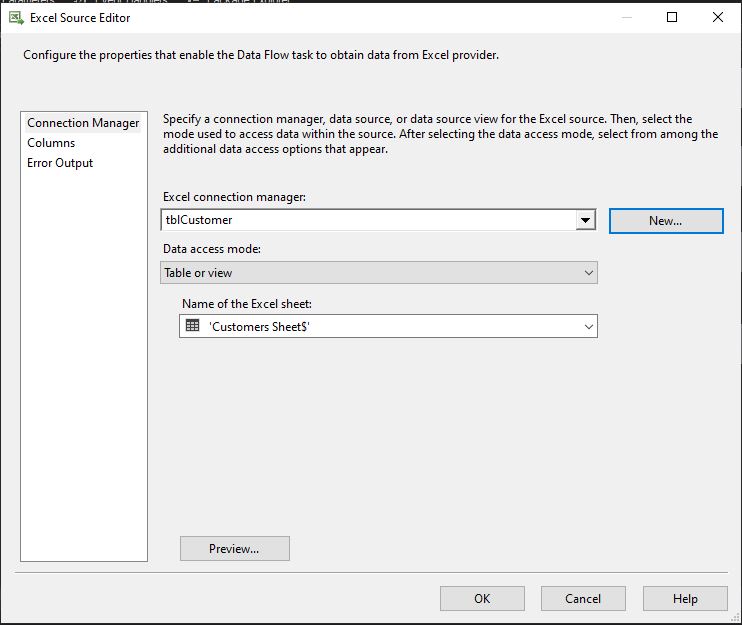
Slika 10. Data Flow “Populate Dimensions” – popunjavanje tabela dimenzija

Set podataka je preuzet u formi jednog excel fajla, gde su tabele bile raspoređene u zasebne listove (engl. *sheets*). Zbog jednostavnosti procesa, svaka tabela je izdvojena i sačuvana kao poseban excel fajl.

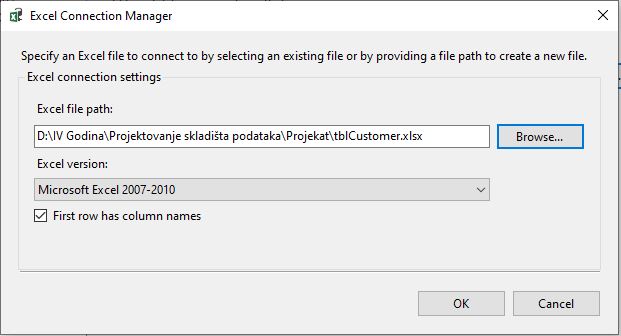
Zbog formata seta podataka, koristi se komponenta za konekciju sa izvorom podataka *Excel Source* i ona takođe zahteva određena podešavanja. (*Slika 11*)

Klikom na opciju *New* otvara se dodatni prozor *Excel Connection Manager* gde se klikom na dugme *Browse* vrši odabir excel fajla koji će biti izvor podataka prilikom punjenja dimnzije *dimCustomer*. (*Slika 12*)

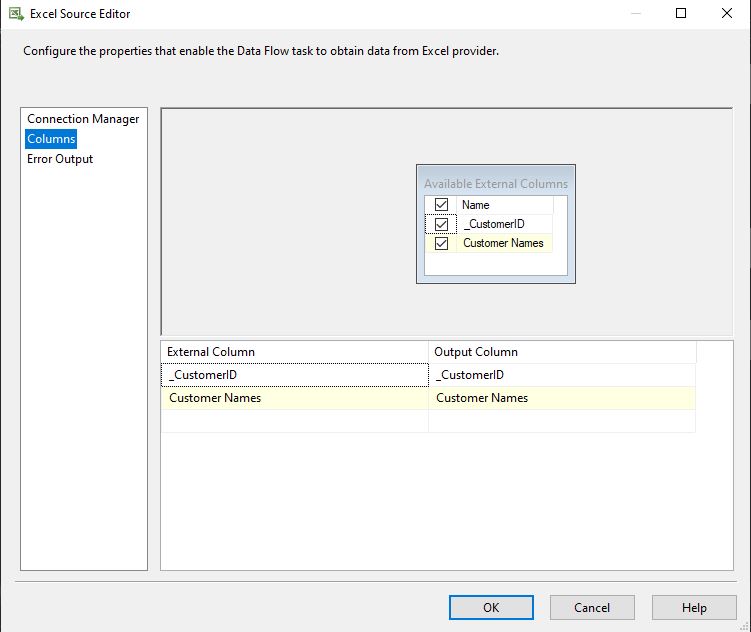
U *Excel Source Editor-*u u kartici *Columns* postoji mogućnost odabira kolona koje želimo da ekstraktujemo. (*Slika 13*)



Slika 11. Excel Source Editor

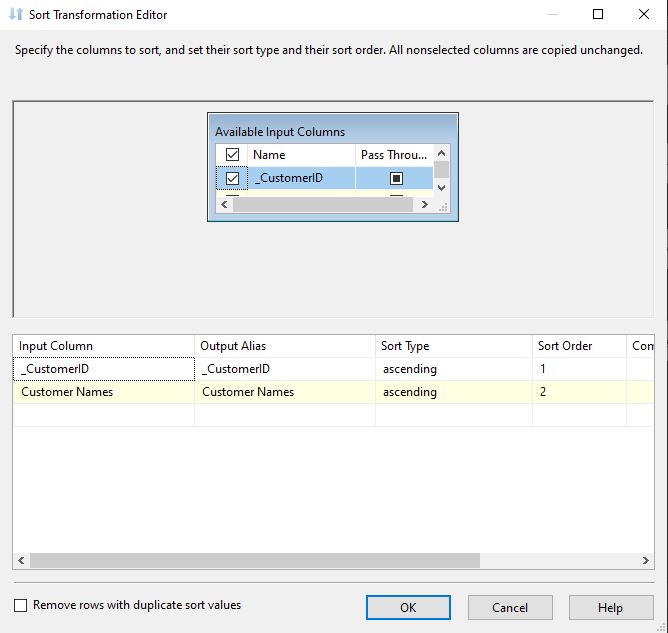


Slika 12. Konekcija sa excel fajlom



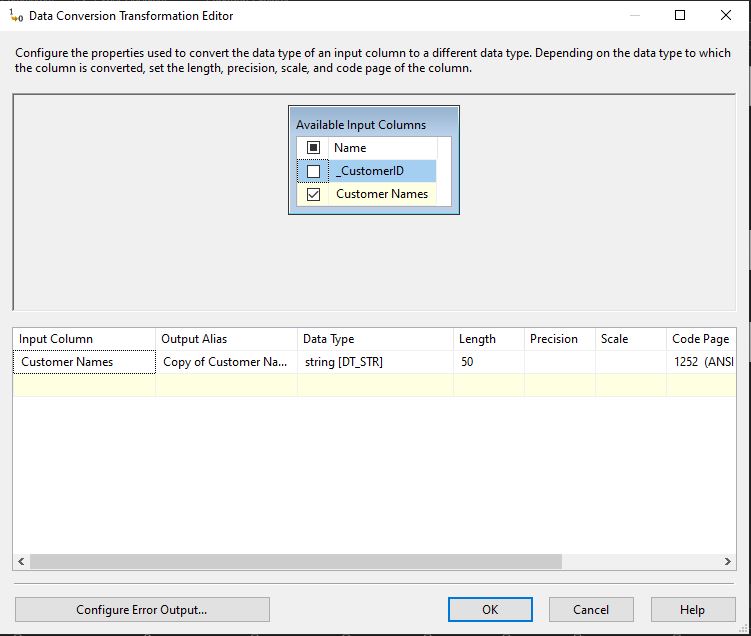
Slika 13. Kartica Columns

Nakon što je konekcija sa izvorom uspostavljena i potrebni podaci su ekstraktovani, neophodno je izvršiti njihovo sortiranje upotrebom komponente *Sort*. Podaci mogu biti sortirani rastuće ili opadajuće odabirom odgovarajuće opcije u polju *Sort Type.* (*Slika 14*)



Slika 14. Sortiranje podataka

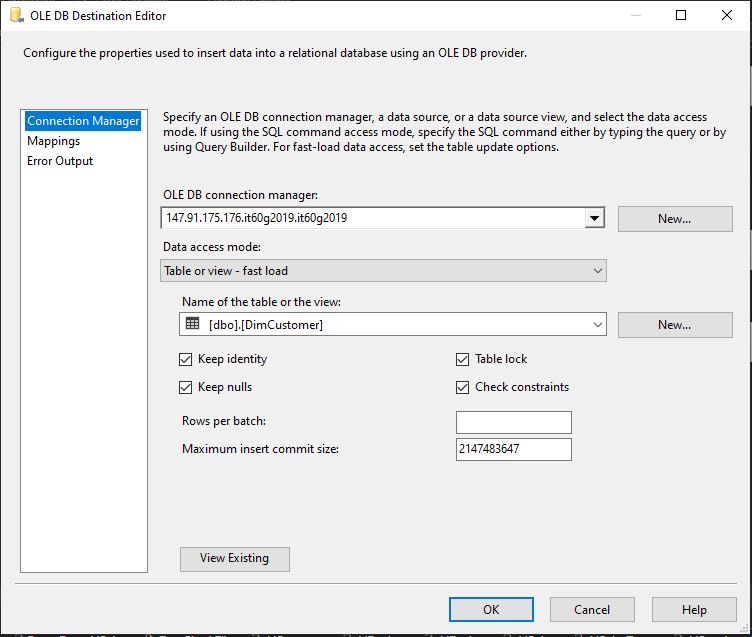
Osim sortiranja, potrebno je izvršiti i konverziju podataka. Transofrmacija podataka se vrši kako bi se ekstraktovani podaci prilagodili formatu podataka u kojem se podaci čuvaju u okviru skladišta podataka. Kada je transformacija izvršena, kreira se kopija transformisanog atributa u odgovarajućem formatu. (*Slika 15*)



Slika 15. Konverzija podataka

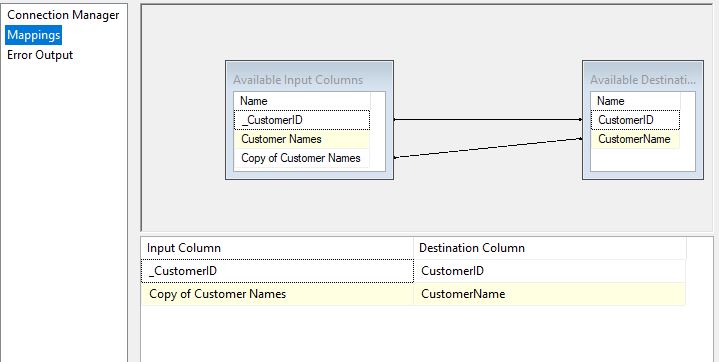
Kada su podaci pripremljeni i prilagođeni, spremni su za punjenje u odgovarajuću tabelu dimenzija. Pomoću komponente *OLE DB Destination* koja je u ovom slučaju preimenovana u “DimCustomer” (*Slika 8*), uspostavlja se konekcija sa tabelom dimenzija koju treba da napunimo podacima ekstraktovanim iz excel fajla.

U *OLE DB Destination Editor-*u se unosi konekcioni string za server na kom su kreirane tabele dimenzija, a zatim se u padajućoj listi bira odgovarajuća tabela dimenzija. Važno je čekirati opciju *Keep Identity* ukoliko je za primarne ključeve korišćeno *Identity* svojstvo kolone. (*Slika 16*)



Slika 16. OLE DB Destination Editor - DimCustomer

U kartici *Mappings* potrebno je mapirati kolone iz izvornih fajlova, sa odgovarajućim kolonama u tabeli dimenzija. Ukoliko je izvršena transformacija podataka i postoji kopija konvertovanih atributa, onda se te kopije koriste prilikom mapiranja umesto originalnih kolona iz izvora. (*Slika 17*)



Slika 17. Kartica Mappings – Mapiranje kolona

Isti proces se ponavlja i za preostale dimenzije.

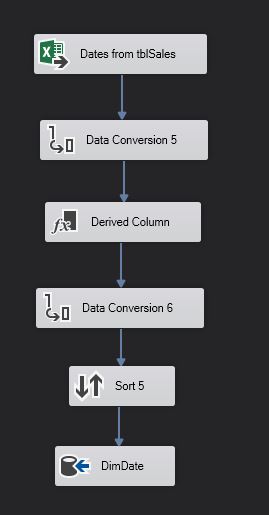
## **Punjenje tabele vremenske dimenzije**

Kada je u pitanju vremenska dimenzija (*dimDate*), postupak za njeno punjenje se nalazi u istom *Data Flow-*u kao i za ostale dimenzije, ali su korišćene različite komponente. (*Slika 18*)

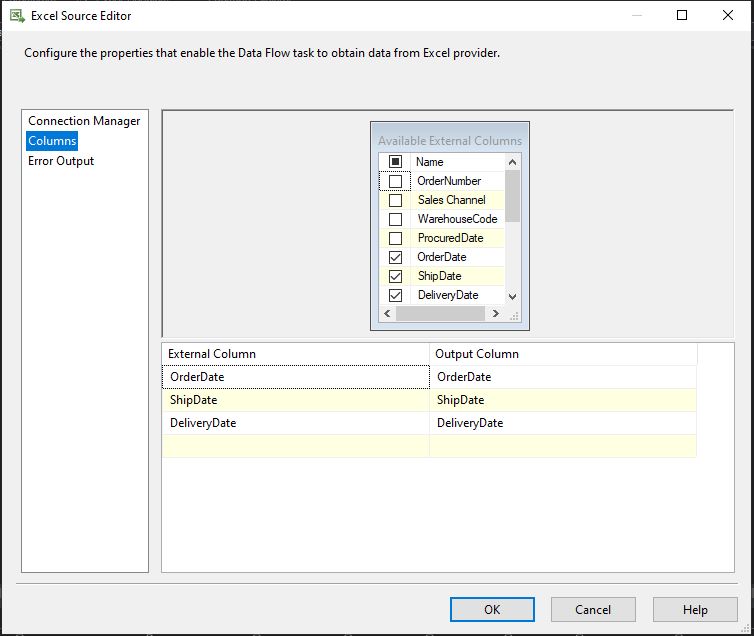
Kao izvor podataka, upotrebljen je excel fajl *tblSales* koji sadrži tabelu od koje je kreirana tabela činjenica. Ovaj put, prosleđene su samo selektovane kolone, potrebne za realizovanje željene granularnosti. (*Slika 19*)

Nakon što su potrebne kolone ekstraktovane iz izvornog fajla, neophodno je izvršiti konverziju kolona u tip podatka *date.* (*Slika 20*)

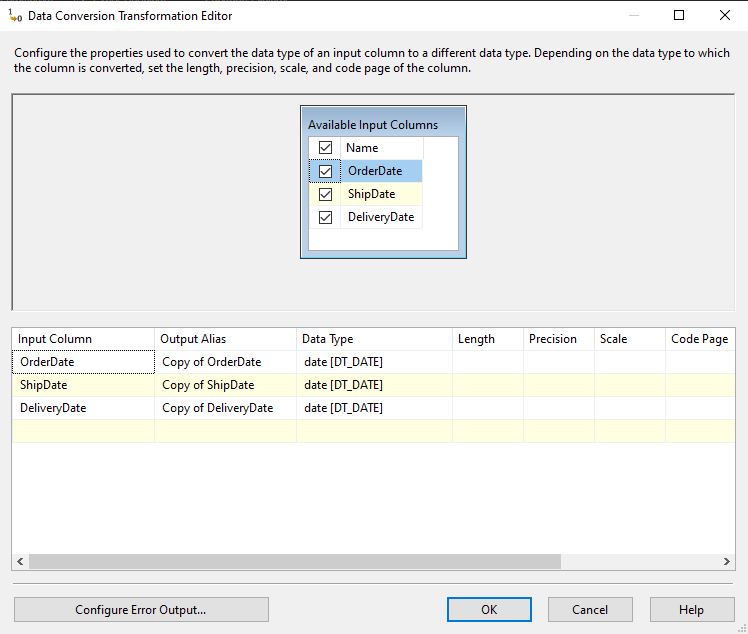
Kako bi se postigla željena granularnost, iz selektovanih datuma (*OrderDate, ShipDate, DeliveryDate*) izdvaja se mesec i godina porudzbine. Izdvajanje je postignuto uz pomoć komponente *Derived Column*, pri čemu je izabrana opcija da se rezultat izdvajanja smesti u novokreiranu kolonu. Takođe zbog bolje preglednosti i razumevanja, umesto broja meseca, piše skraćeni naziv što je omogućeno upotrebom ugrađene funkcije *SUBSTRING*. Godina je izdvojena iz datuma porudzbine upotrebom ugrađene funkcije *YEAR.* (*Slika 21*)



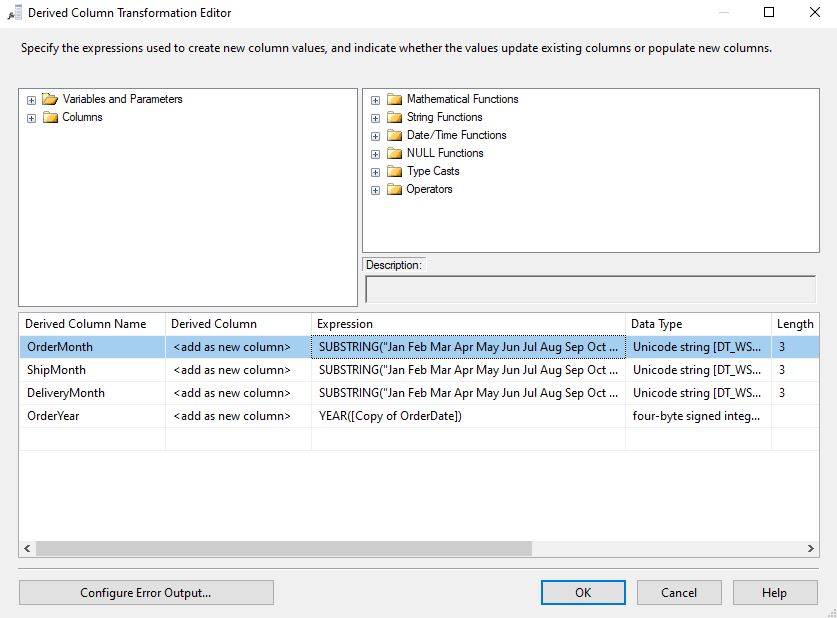
Slika 18. Postupak punjenja vremenske dimenzije dimDate



Slika 19. Prosleđene kolone iz excel fajla tblSales



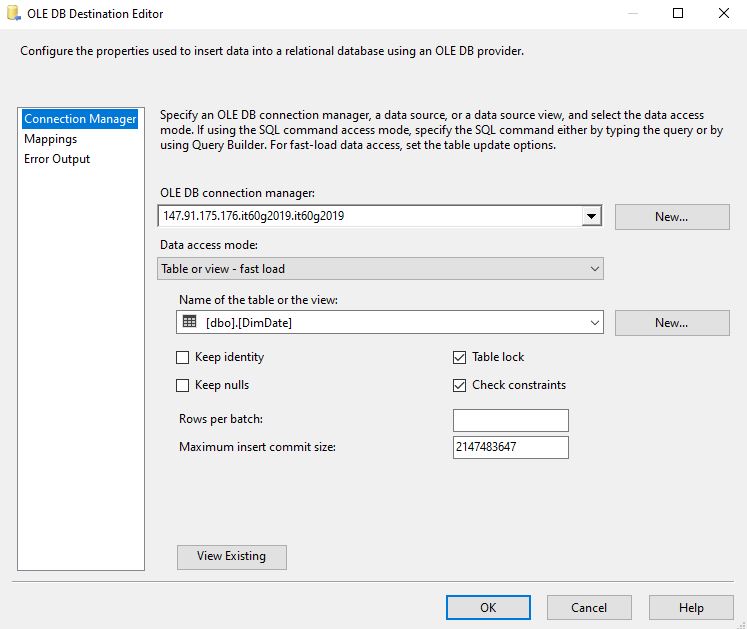
Slika 20. Konverzija selektovanih kolona u odgovarajući tip podatka



Slika 21. Komponenta Derived Column – Izdvajanje meseca i godine

Novokreirane kolone za naziv meseca porudzbine, isporuke i prispeća koje su rezultat komponente *Derived Column* treba konvertovati u tip podatka *string,* a zatim izvršiti sortiranje podataka kao i kod prethodnih punjenja.

Što se tiče destinacije, odnosno tabele vremenske dimenzije, takođe se koristi komponenta *OLE DB Destination,* unosi se konekcioni string, ali ovaj put opcija *Keep identity* nije čekirana jer je primarni ključ u tabeli dimenzije *dimDate* datum porudzbine (*OrderDate*). Zbog toga je prilikom sortiranja čekirana opcija da se eliminišu duplikati. (*Slika 22*)



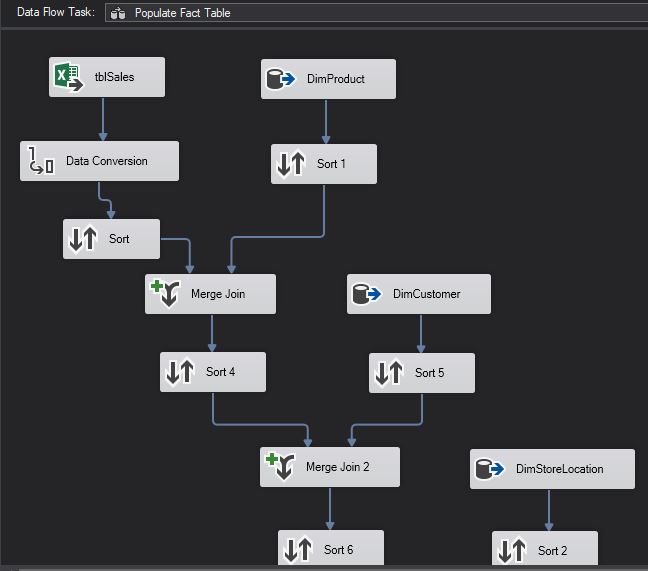
Slika 22. OLE DB Destination Editor za tabelu dimenzije dimDate

Završni korak je mapiranje novokreiranih kopija kolona usled konverzije podataka i izvođenja novih kolona sa odgovarajućim kolonama u okviru tabele dimenzije *dimDate.* Postupak je isti kao i kod prethodnih dimenzija.

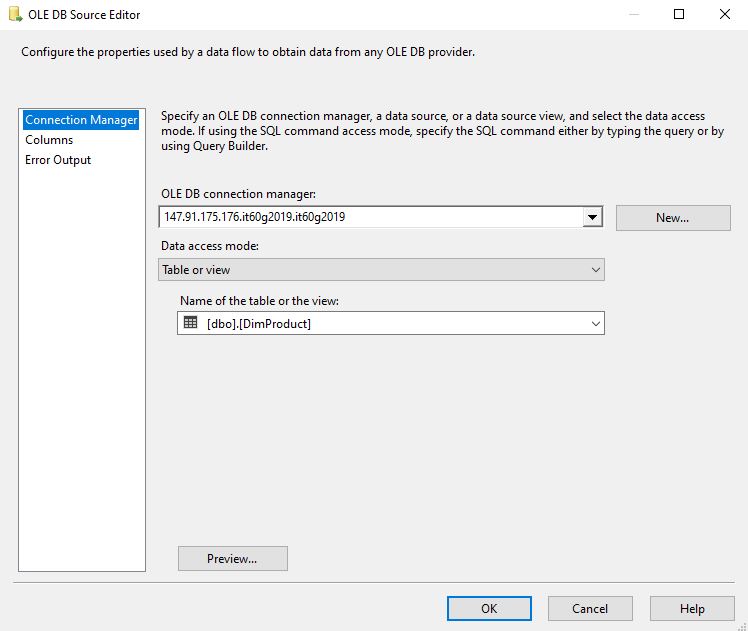
## **Punjenje tabele činjenica**

Iako se tabela činjenica kreira istovremeno kada i dimenzije, da bi se ona popunila neophodan je poseban *Data Flow* koji se inicira tek nakon što su popunjene sve tabele dimenzijai kao što je već prikazano ranije u tekstu on je preimenovan u “Populate Fact Table”. (*Slika 8*)

Proces punjenja podataka u tabelu činjenica inicira se upotrebom excel fajla *tblSales,* odnosno izvorne tabele *Sales* od koje je kreirana tabela činjenica i prosleđuju se sve kolone. Tabela činjenica se popunjava na osnovu prethodno popunjenih tabela dimenzija, pa se zato koristi komponenta *OLE DB Source.* Komponenta je preimenovana u “DimProduct” jer su sada izvor podataka tabele dimenzija kreirane na serveru, te se definiše konekcija na isti i bira se odgovarajuća tabela. (*Slika 23-24*)

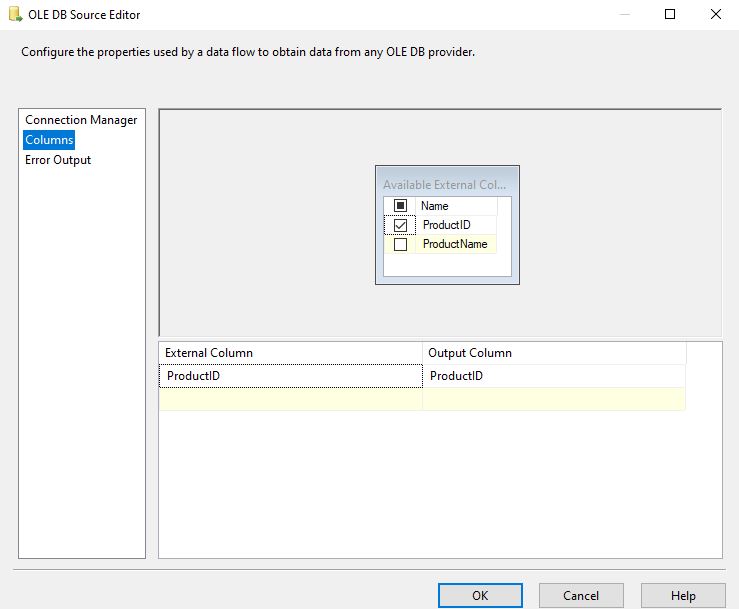


Slika 23. Punjenje tabele činjenica 1



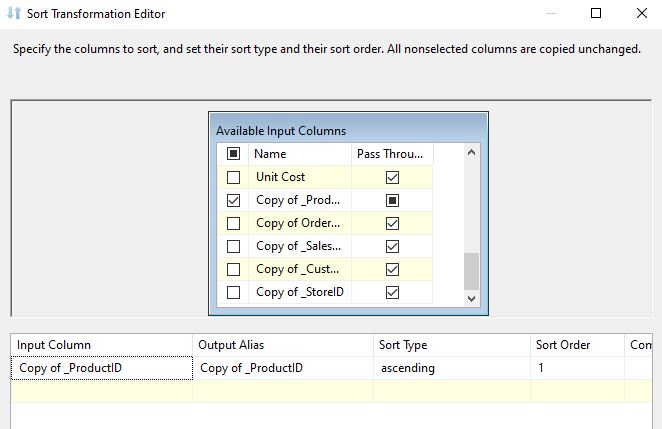
Slika 24. Konekcija na tabelu dimenzija dimProduct za prvi spoj

Ovaj put se iz tabela dimenzija preuzima samo primarni ključ radi vršenja spoja upotrebom komponente *Merge join*. (*Slika 25*)



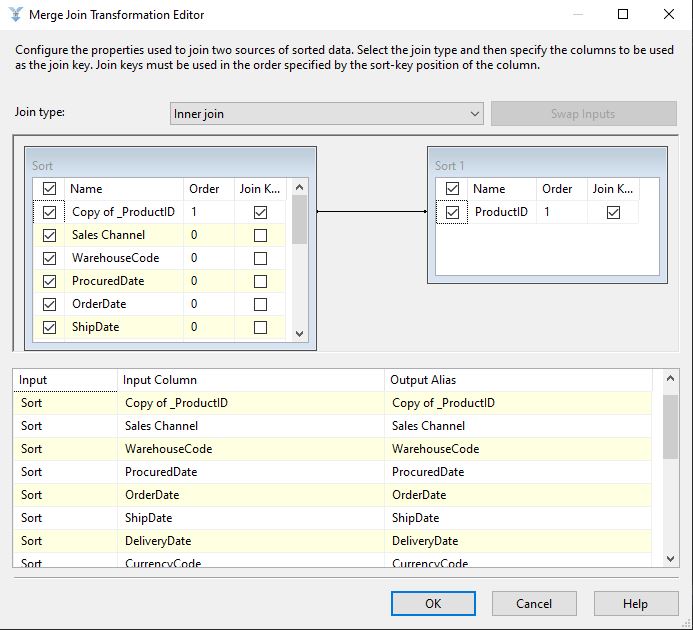
Slika 25. Prosleđivanje ključa iz tabele dimenzija kako bi se po ključu vršio spoj

Kako bi bilo moguće izvršiti spoj tabela po ključevima u komponenti *Merge join,* prvo je potrebno izvršiti sortiranje za oba izvora, ali i sortiranje se ovaj put vrši samo po ključevima po kojima je predviđeno spajanje, ostale potrebne kolone se samo prosleđuju čekiranjem opcije *Pass Through.* (*Slika 26*)



Slika 26. Sortiranje po ključu za spoj

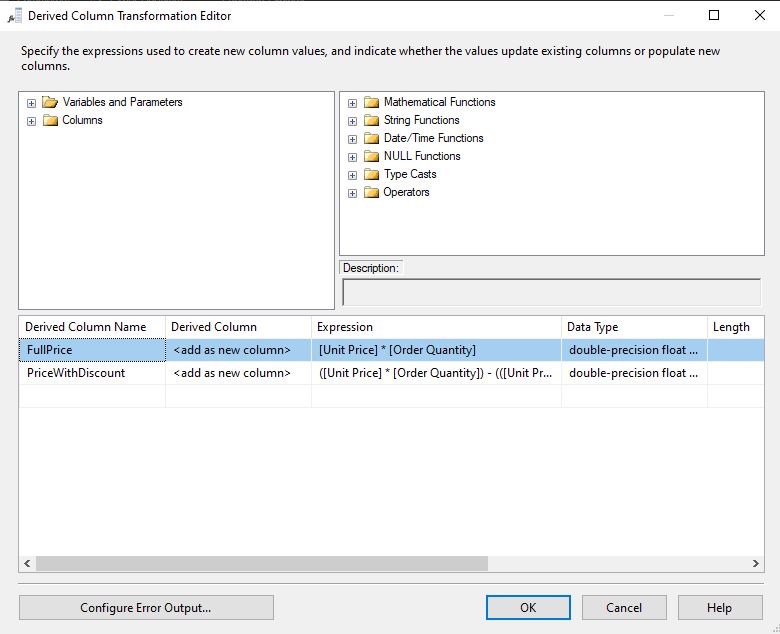
Kada je sortiranje po ključu za oba izvora, *Excel Source* i *OLE DB Source,* završeno, tada je moguće izvršiti spajanje pomoću komponente *Merge join.* Izabran je tip unutrašnjeg spoja, ključevi su automatski mapirani i neophodne kolone se idalje prosleđuju dalje. (*Slika 27*)



Slika 27. Postupak spajanja tabela u komponenti Merge join

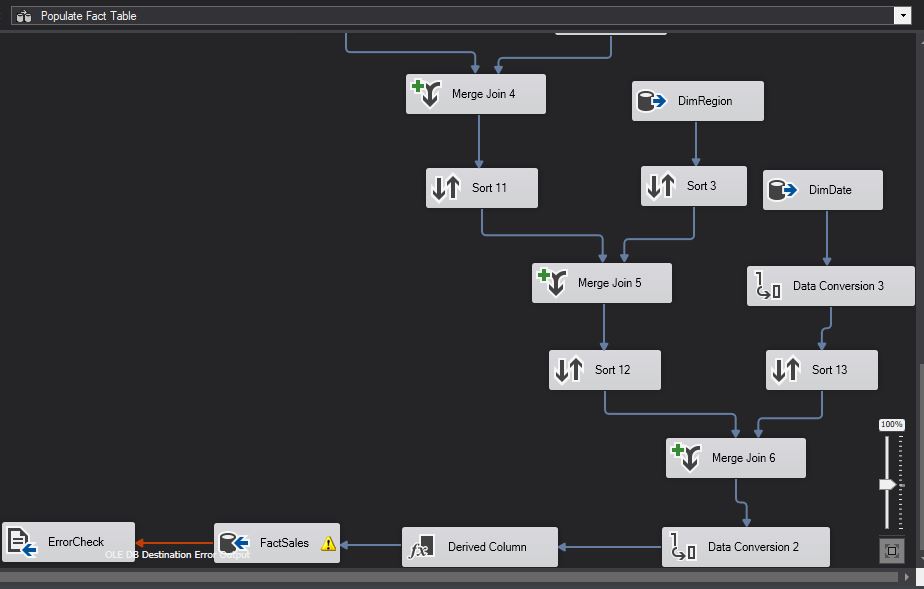
Na prethodnim slikama prikazana je logika spajanja tabela za potrebe punjenja tabele činjenica. Postupak se nastavlja na isti način, spajanjem svake od tabela dimenzija i preuzimanjem njihovog ključa kako bi se izvršio spoj, dok se ostale kolone, svojstvene tabeli činjenica, preuzimaju iz izvornog excel fajla *tblSales* i prosleđuju se do kraja samog procesa punjenja.

Na samom kraju procesa, uz pomoć komponente *Derived Column,* kreiraju se dodatne dve kolone, ranije opisane u odeljku sa podnaslovom “Specifikacija zahtevanih mera”, *FullPrice* i *PriceWithDiscount.* (*Slika 28-29*)

**

Slika 28.Drived Column - Kreiranje zahtevanih mera

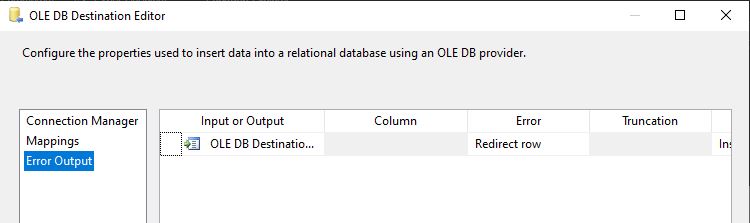
Na kraju procesa spajanja, definiše se konekcija na tabelu činjenica pomoću komponente *OLE DB Destination* koja je preimenovana u “FactSales” na isti način kao i što je prikazano i na prethodnim slikama, što važi i za mapiranje. (*Slika 29*)



Slika 29. Prikaz završnih koraka u procesu popunjavanja tabele činjenica

Može se primetiti na prethodnoj slici da postoji još jedna komponenta koja do sad nije korišćena. U pitanju je *Flat File Destination* komponenta preimenovana u “ErrorCheck” koja ima samo jednu svrhu, kao što sam naziv kaže, da proveri moguće greške prilikom upisa podataka u tabelu činjenica. Razlog dodavanja ove komponente je zbog prevelike količine podataka, a samim tim i veće šanse da se desi propust. Ukoliko prilikom procesa punjenja dođe do neke greške koja će prekinuti izazvati prekid procesa, podaci će se upisati u tekstualni fajl sa kojim je izvršena konekcija kroz komponentu *Flat File Destination.*

Redirekcija upisa u slučaju greške je omogućena u kartici *Error Output* u okviru *OLE DB Destination Editor-*a “FactSales” izmenom vrednosti polja *Error* u “Redirect row”*.* (*Slika 30*)

**

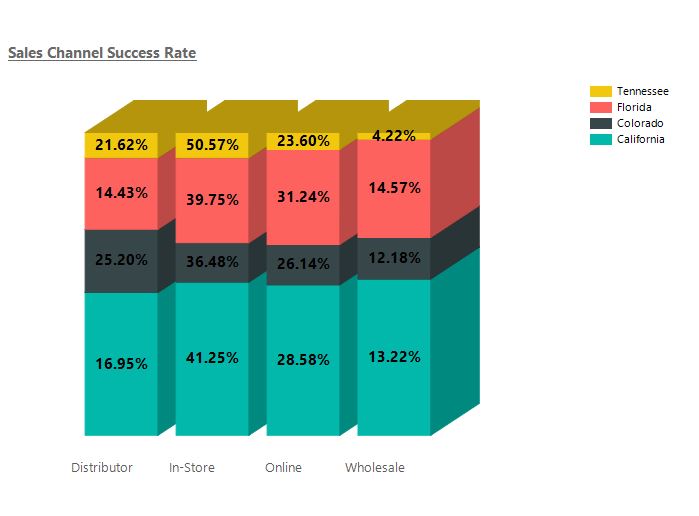
Slika 30. Redirekcija upisa redova prilikom greške

Ovim korakom se završava ETL proces i skladište podataka je spremno za kreiranje ad-hock zahteva.

# **Izveštaji**

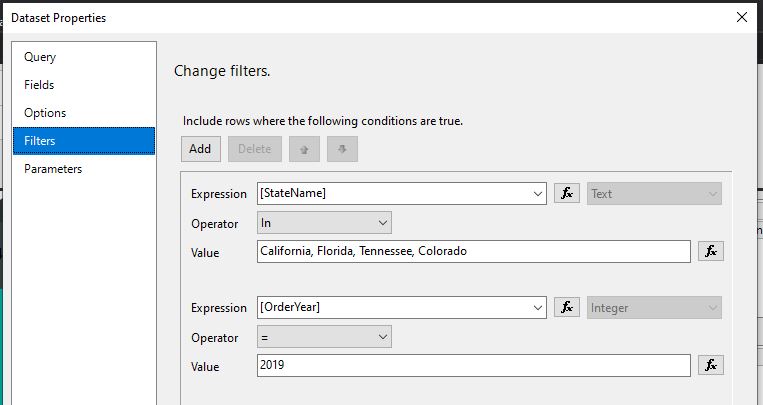
## **Prvi izveštaj**

Prvi izveštaj prikazuje uspešnost prodaje preko svakog kanala prodaje, izražene procentualno, na teritorijama navedene četiri države za 2019. godinu. Grupe serija su države, a grupe kategorija su kanali prodaje. Uspešnost se meri na osnovu ukupnog iznosa cena sa uračunatim popustom. *Data set* omogućava i merenje u odnosu na cenu bez uračunatog popusta. (*Slika 31*)



Slika 31. Sales Channel Success report

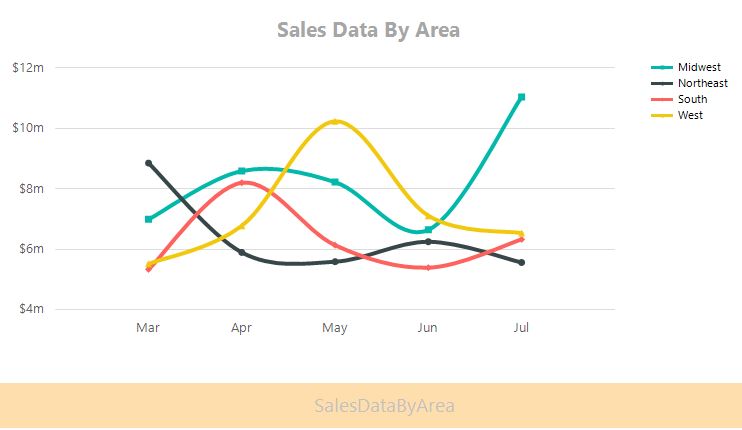
Filtriranje je izvršeno na sledeći način. (*Slika 32*)



Slika 32. Filtriranje u prvom izveštaju

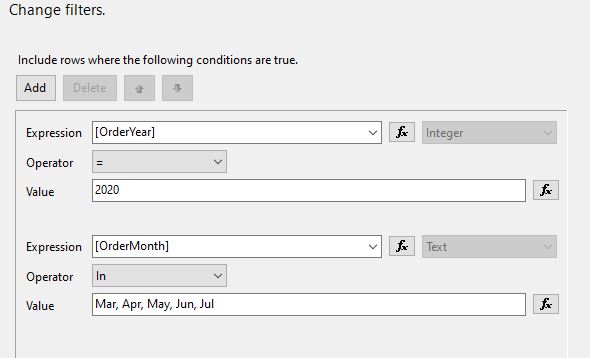
## **Drugi izveštaj**

Drugi izveštaj pokazuje linijsko kretanje prihoda kroz 5 meseci u 2020. godini, počevši od marta meseca, u posmatranim regijama. Cilj izveštaja je da prikaže stabilnost ukupnog prihoda kroz određeni period. Rezultat izveštaja je detaljno opisan u odeljku sa podnaslovom „Specifikacija zahteva korisnika“ pod tačkom broj 2. Izveštaj je prikazan na sledećoj slici. (*Slika 33*)



Slika 33. Sales data by area report

Izvršeno je filtriranje seta podataka za posmatrani izveštaj kako bi se dobile tražene informacije. (*Slika 34*)



Slika 34. Filtriranje za drugi izveštaj

## **Treći izveštaj**

Treći izveštaj predstavlja pregled broja proizvoda koji su prodati na popustu većem od 10%, koji tim prodaje je vršio tu prodaju, prikaz pune cene i cene sa uračunatim popustom kao i popust izražen procentualno. Za određene vrednosti se računa i ukupan rezultat. U izveštaj su dodati heder koji sadrži naslov izveštaja i futer koji prikazuje trenutnu stranicu od ukupnog broja stranica. (*Slika 35-36*)

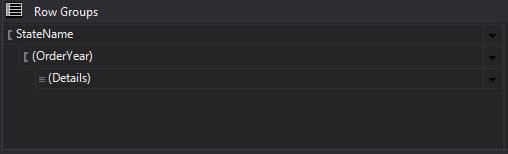


Slika 35. Products sold with discount report



Slika 36. Prošireni prikaz izveštaja

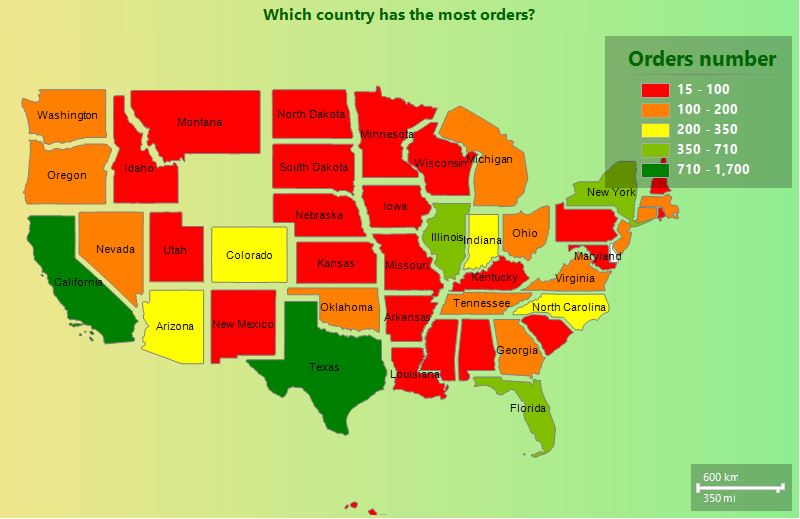
Izvršeno je grupisanje prvo po godinama, a zatim i po državama u kojima su se vršile posmatrane prodaje. (*Slika 37*)



Slika 37. Grupisanje

## **Četvrti izveštaj**

Ovaj izveštaj vizuelno prikazuje na mapi u kojim državama se najviše porudzbina izvršava, a u kojim najmanje. Države sa najmanje porudžbina su označene crvenom bojom, a one sa najviše su označene zelenom. *Data set* omogućava varijacije upita gde umesto prebrojavanja porudzbina može da se posmatra naseljenost država, ili srednji prihod na nivou istih. (*Slika 38*)



Slika 38. MapInfo report

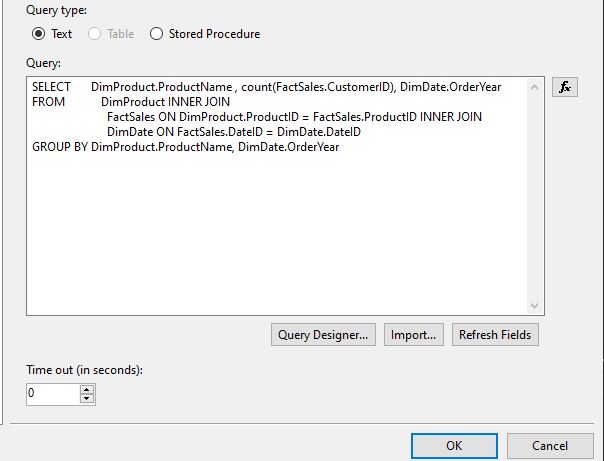
## **Peti izveštaj**

Peti izveštaj prikazuje top 5 najprodavanijih proizvoda za svaku godinu. (*Slika 39*)



Slika 39. Top5 Best selling Products report

Kolona *Order number* je zapravo nastala prebrojavanjem jedinstvenih identifikacionih oznaka kupaca spram kategorije proizvoda koje su kupovali. Grupisanje je izvršeno još na samom početku kreiranja *data set-*a. (*Slika 40*)



Slika 40. Grupisanje u DataSet properties

# **Zaključak**

Proces planiranja, modelovanja i razvoja skladišta podataka za potrebe izrade projekta je uporedo praćen dokumentacijom u okviru koje su detaljno opisani postupci projektovanja i pristup samoj izradi. U zavisnosti od potreba krajnjeg korisnika, menja se i pogled na podatke. Iz tog razloga, skladište podataka je uvek podložno mogućim izmenama i prilagođavanju. Na osnovu jednostavnog primera koji je obrađen u okviru ovog projekta, može se zaključiti koliko je proces projektovanja skladišta podataka zahtevan i kompleksan proces. Razumevanje tematike, oblasti poslovanja, pravila poslovanja, potreba korisnika i mnogih drugih segmenata je od ključne važnosti za krajnji rezultat i funkcionalnost samog DW sistema.