



20. konferenca Dnevi slovenske informatike

SESTAVA GALAKSIJA SHEME PODATKOVNEGA SKLADIŠČA ZA NAMEN SPREMLJANJA ZALOG, NABAVE, PROIZVODNJE IN PRODAJE



Igor Makovec

15. 04. 2013



Poslovna inteligenca (*angl. business intelligence, BI*)

Definicija

Poslovna inteligenca je proces, tehnologije in orodja potrebna za pretvarjanje podatkov v informacije, informacije v znanje in znanje v plane, ki privedejo do profitabilnih poslovnih akcij. Poslovna inteligenca obsega podatkovno skladiščenje, poslovno analitska orodja in upravljnje znanja.

(Data Warehouse Institute)





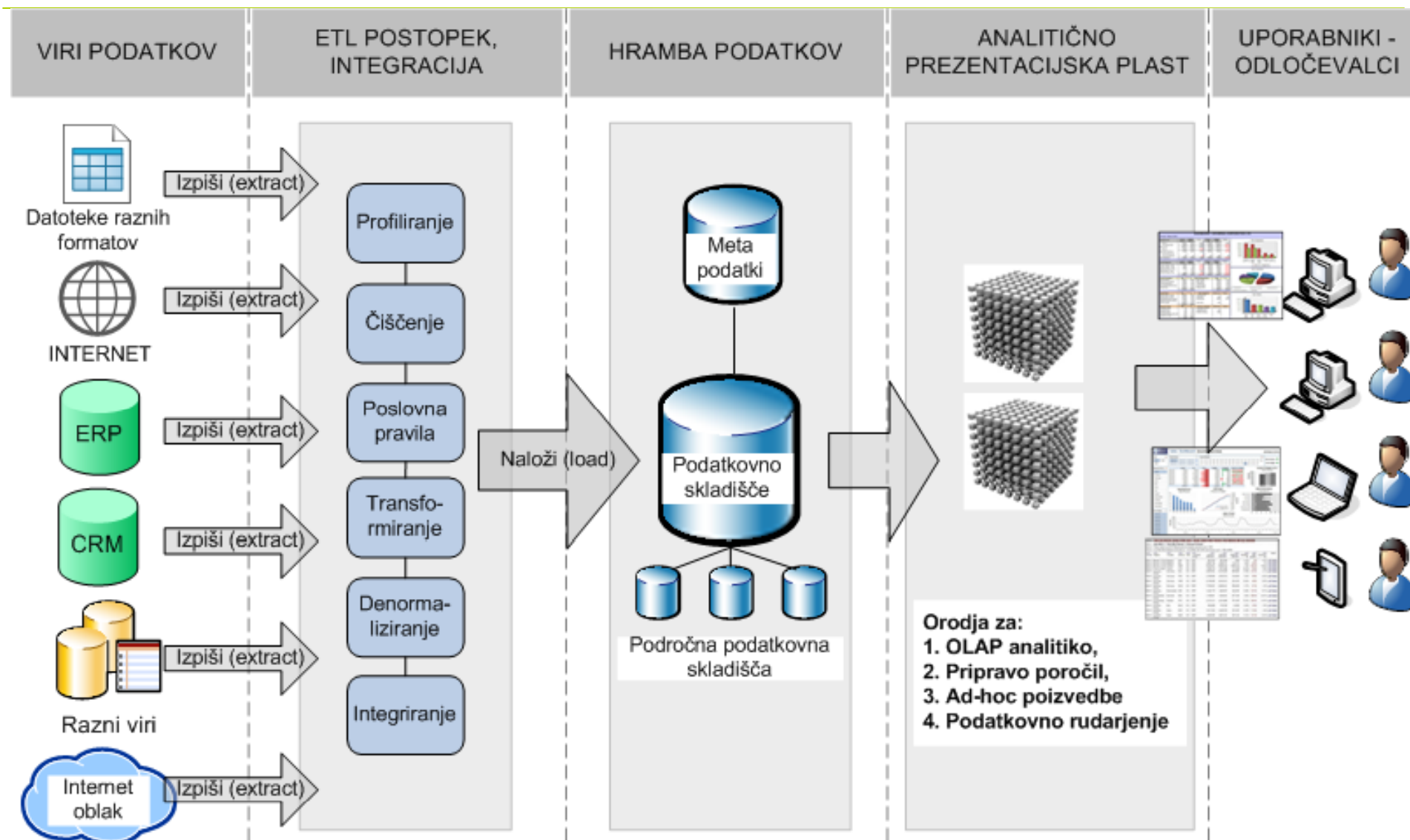
Zakaj uvesti BI:



1. Pridobljene informacije omogočajo boljše odločanje;
2. Informacije so hitre in lahko pravočasno sprejmemo in izvedemo aktivnosti;
3. Omogočena je integracija podatkov iz različnih virov vključujoč podatke iz transakcijskih sistemov, raznih podatkovnih baz, datotek različnih formatov zapisa in podatkov iz spleta;
4. Informacije lahko prikažemo grafično, kar je bolj nazorno;
5. Kreiranje nadzornih plošč je enostavno in ga lahko izvedejo kar uporabniki sami, brez posredovanja IT oddelka;
6. Uporabniški vmesniki so enostavni, tudi tipa povleci-spusti in omogočajo nastavitve za kompleksne analize na enostaven intuitiven način;
7. Omogočene so analize ogromno podatkov v realnem času;
8. Razkrite informacije lahko delimo s sodelavci preko internetnega omrežja preko vseh kanalov: email, standardna poročila, mobilne vsebine, socialna omrežja;
9. Informacije so dostopne na mobilnih napravah in v vsakem trenutku.



Infrastruktura BI





Sestava podatkovnega skladišča (angl. data warehouse, DW)

Golfarelli in Rizzi (1998) sta ugotavljala neobstoječe in konsistentne metodologije za načrtovanje podatkovnega skladišča in sta predlagala 6 faz razvoja DW:

Št.	Korak	Vhod	Izhod	Udeleženci
1	Analiza informacijskega sistema	Obstoječa dokumentacija	Shema DW	DW designer, IT manager
2	Zbiranje zahtev uporabnikov	(shema DW)	Seznam omogoča identificirati dejstva in dimenzije	DW designer, končni upravniki
3	Konceptualni design	Shema DW	Dimenzijska shema	DW designer
4	Izboljšava in validacija designa	Dimenzijska shema	Dimenzijska shema	DW designer, končni upravniki
5	Logični design DW	Dimenzijska shema	Logična DW shema	DW designer
6	Fizični design DW	Logična DW shema	Fizična DW shema	DW designer



3. Konceptualni design DW

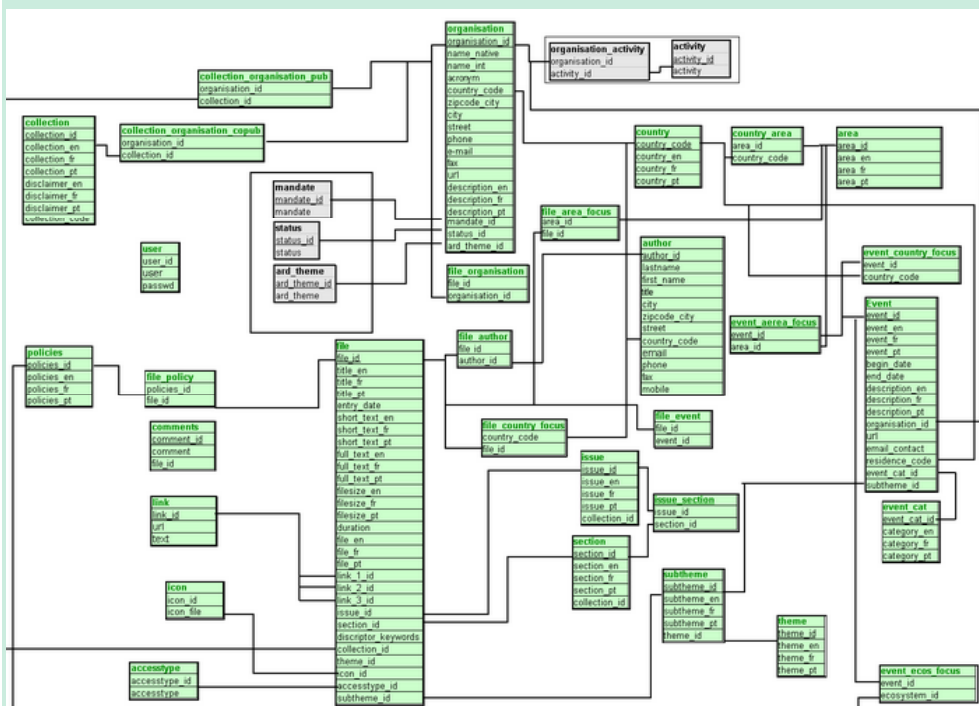
Razlogi zakaj kreirati DW se skrivajo v razlikah med ERP transakcijskim sistemom OLTP (angl. online transaction processing) in OLAP (angl. online analytical processing) sistemom.

OLTP	OLAP
Podatki so shranjeni v ogromnem številu normaliziranih tabel. Organizacija podatkov ne ustreza analitikom. Izvedba poizvedb je zapletena.	Podatki so shranjeni v podatkovnem skladišču. Večdimenzionalni model. Naravna podpora za izračune agregatov (vsote..), obravnavanje časa in druga izračunavanja.
Uporabniki: vsi zaposleni	Uporabniki: običajno odločevalci, analitiki
Operativni podatki.	Podatki lahko iz več virov, konsolidirani.
Ažuriranje podatkov trenutno.	Ažuriranje periodično – odvisno od procesa. Običajno enkrat dnevno.
Analitika upočasnjuje sistem – performančni problemi.	Ni obremenjevanja transakcijskega sistema – podatki so v podatkovnem skladišču.
Poročila so pogosto obsežna a neuporabna.	Uporabniki lahko poročila kreirajo sami.



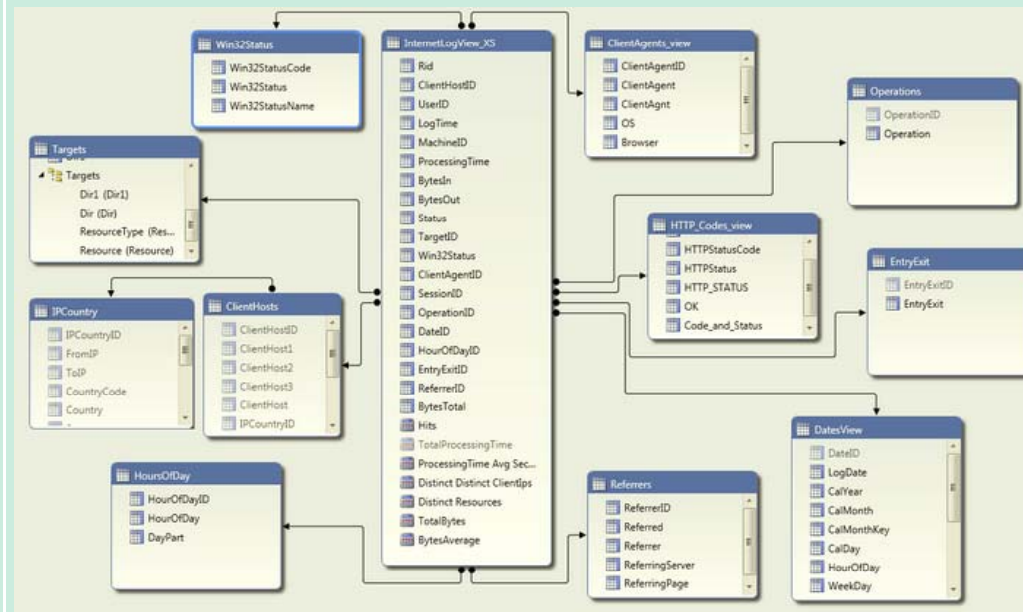
OLTP

Veliko normaliziranih tabel, poizvedbe zaradi veliko povezav komplicirane.
ERP SAP ima več 10.000 tabel.



OLAP

Tabele denormalizirane, poizvedbe enostavne.





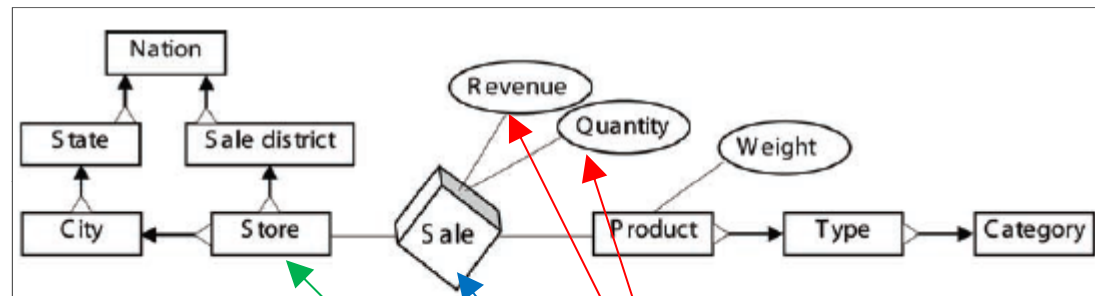
- Konceptualni design DW je ključnega pomena.
 - Rezultat je grafična notacija, ki omogoči hitrejše nadaljne korake.
- Standarda pri tem še nimamo.** Obstaja več predlogov, ki jih lahko strnemo v 4 skupine:

Razširjeni ER modeli	Objektno orientirani modeli	Ontološki modeli	Ad-hoc modeli
<ol style="list-style-type: none"> MER (Multidimensional Entity Relationship Model) MultiDimER starER, SERM (Structured Entity Relationship Model) 	<ol style="list-style-type: none"> YAM (Yet Another Multidimensional Model) UML model, GOLD model 	Niso posebej poimenovani.	<ol style="list-style-type: none"> MAC (Multidimensional Aggregation Cube Data Model) DFM (Dimensional Fact Model) Fact schema model

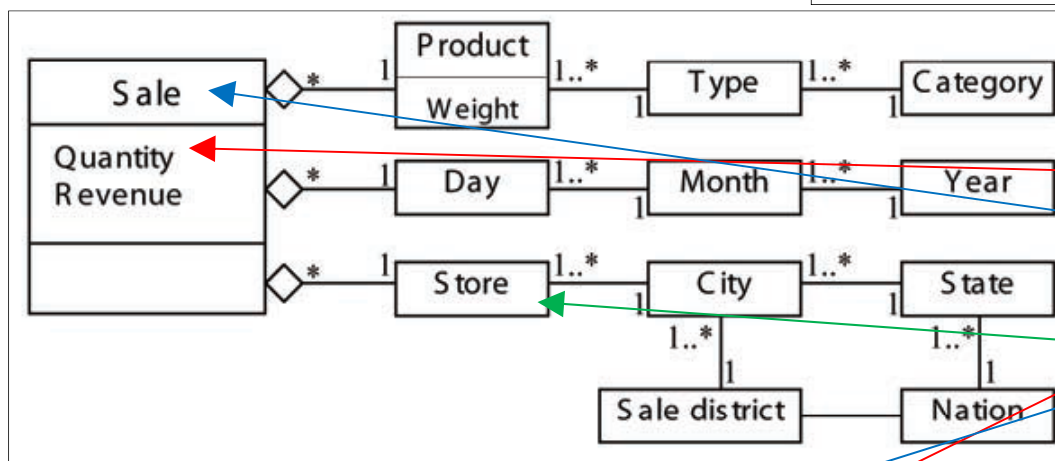


Primeri konceptualnih designov enakega DW:

StarER model



UML model

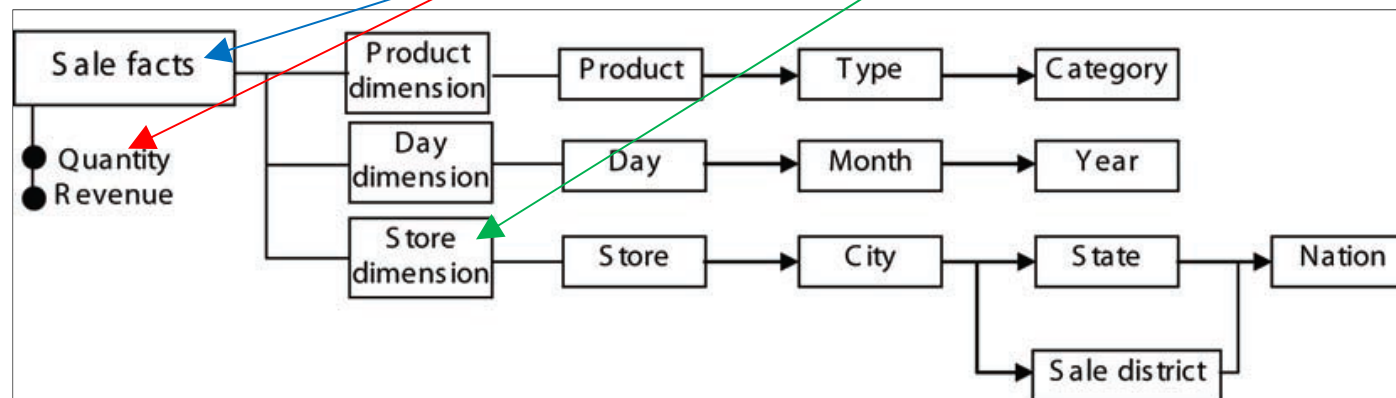


Atributi

Dejstva (fact)

Dimenzije

Fact schema model





5. Logični design DW

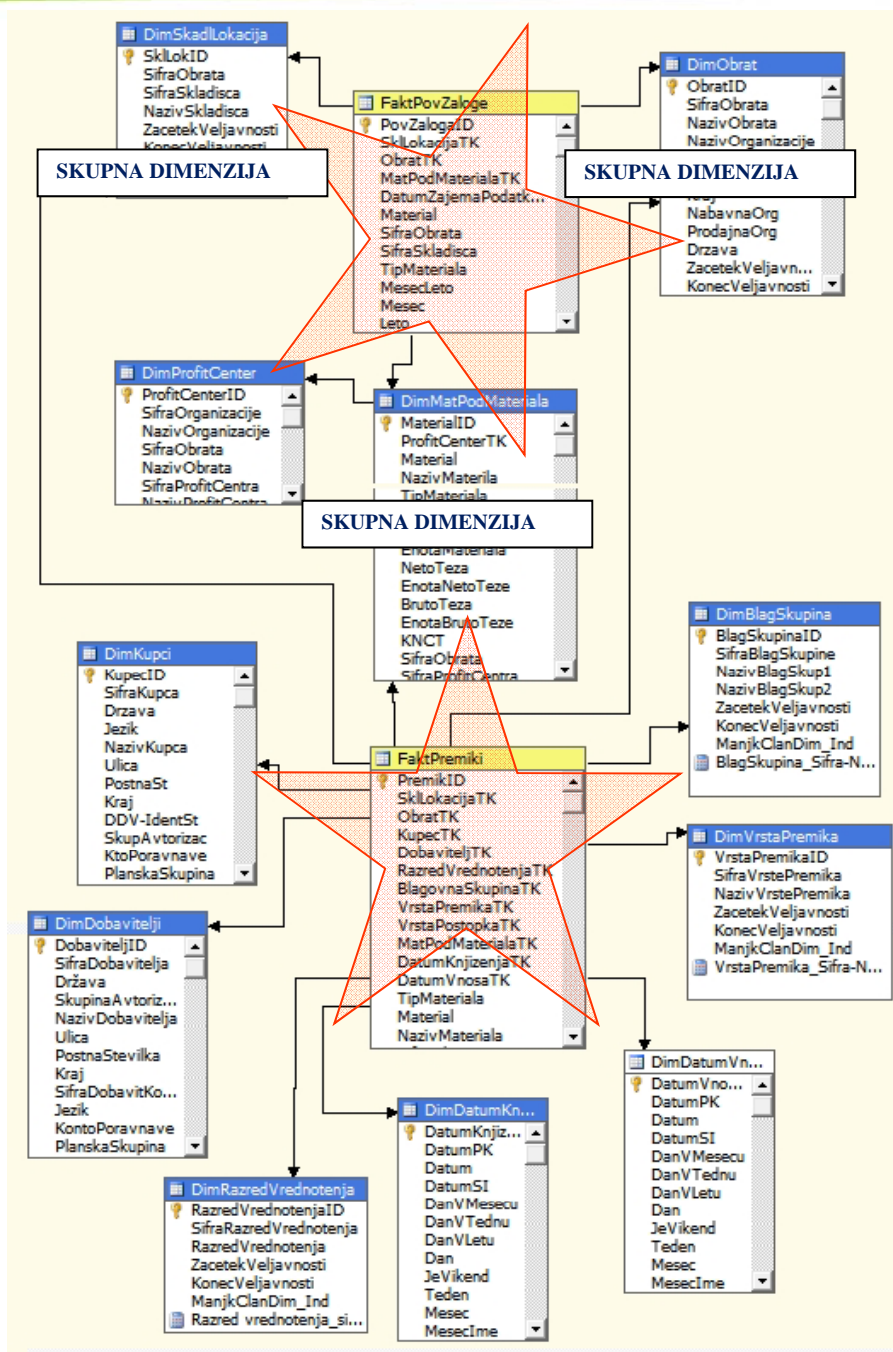
Gre za pretvorbo konceptualne sheme v ciljni podatkovni model bodisi relacijski ali večdimenzijski.

V našem primeru je bil izbran večdimenzijski model zaradi minimalnega števila tabel in hitrosti poizvedb in možnosti uporabe funkcionalnosti OLAP kocke.

Večdimenzijski modeli:

1. ploska shema (*angl. flat schema*),
2. terasasta shema (*angl. terraced schema*),
3. zvezdasta shema (*angl. star schema*),
 - sheme sozvezdja,
 - sheme galaksije
4. shema skupine zvezd (*angl. star cluster schema*),
5. snežinkasta shema (*angl. snowflake schema*),







6. Fizični design DW

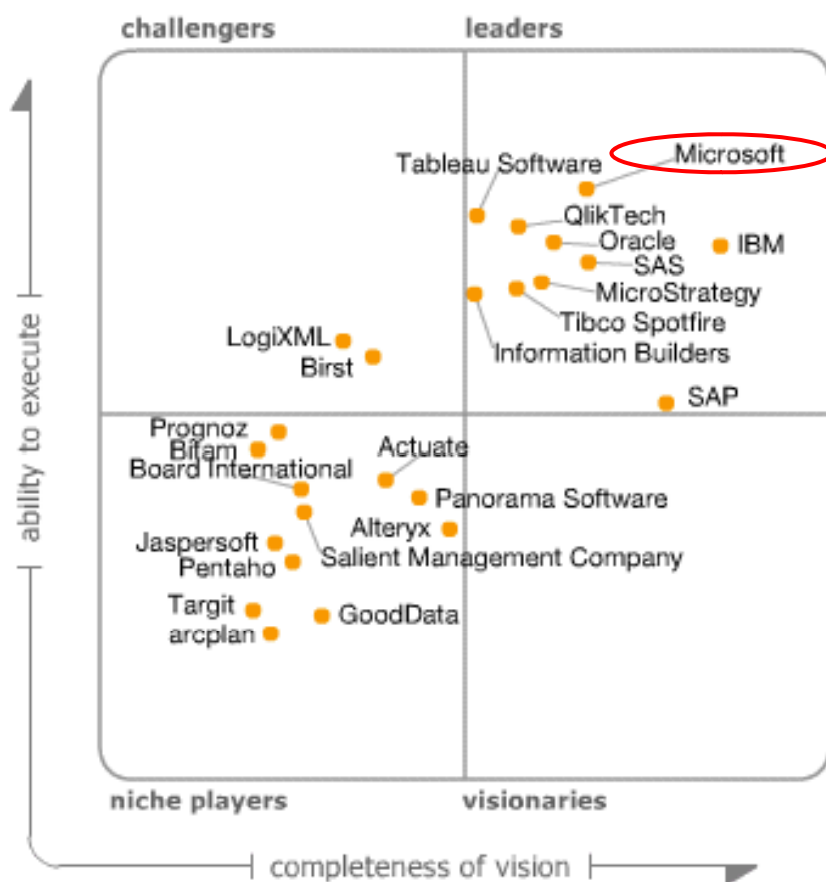
Uporabljeno je bilo prototipiranje.

Življenski cikel prototipa:

1. ugotavljanje informacijskih potreb,
2. razvoj osnovnega prototipa,
3. uporaba in beleženje sprememb,
4. izboljšanje prototipa in vrnitev na točko 3 (če je potrebno),
5. ocena delujočega sistema,
6. izvedba sprememb,
7. instalacija, delovanje in vzdrževanje.



Izbor SW opreme



Slika: Gartnerjev magični kvadrant BI in analitičnih platform

Prednosti po Gartnerju:

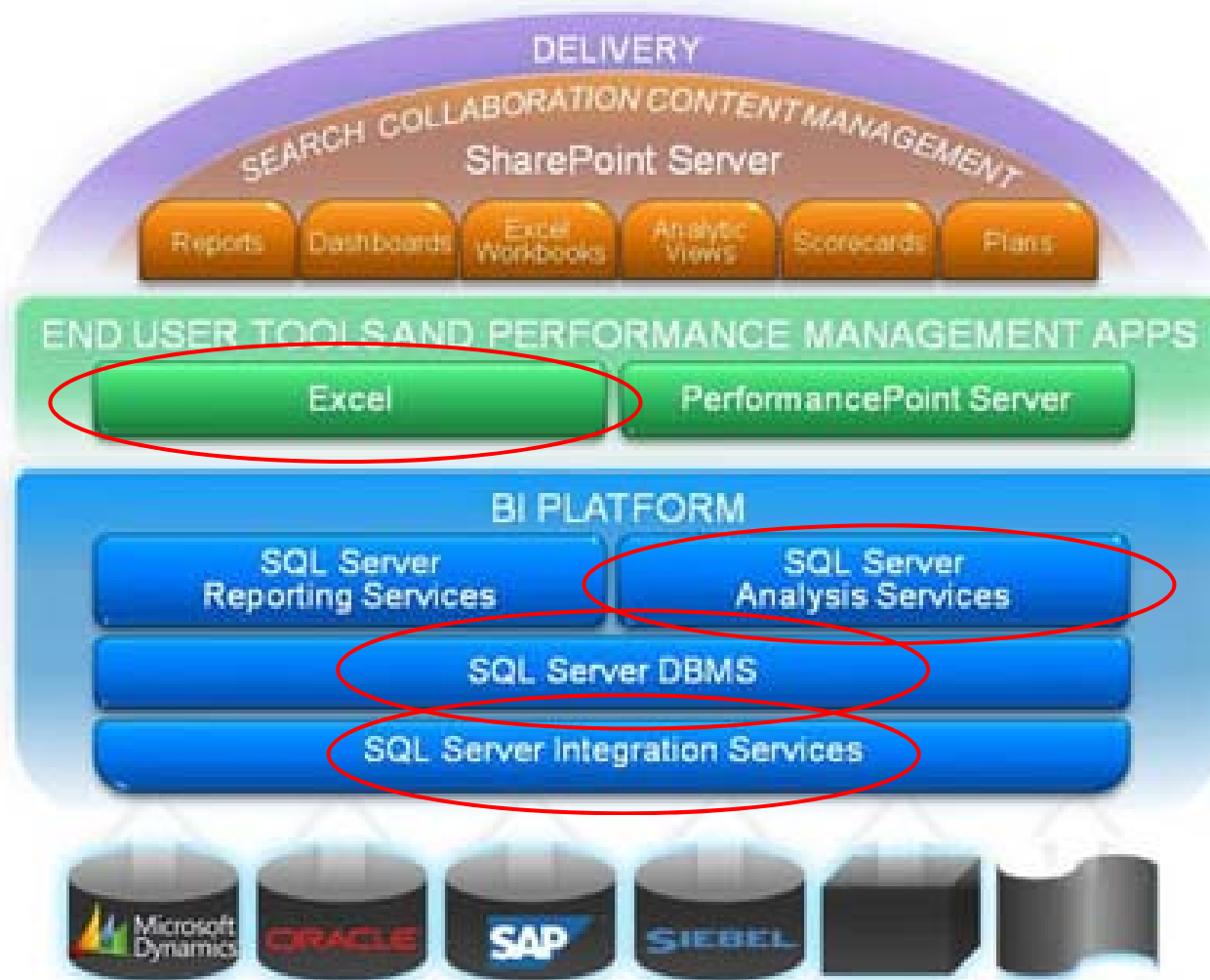
1. nizki licenčni stroški,
2. široka dostopnost znanj,
3. stalna tržna rast,
4. velika večina kupcev uporablja OLAP funkcionalnosti

Slabosti po Gartnerju:

1. kompleksnost zaradi uporabe več produktov, vsak v svojem delu BI,
2. koncept „do-it-yourself“ povzroča težave malim podjetjem,
3. slabša podpora mobilnim napravam,
4. nizki licenčni stroški se relativizirajo, saj stroški implementacije in razvoja presežejo 75% vseh stroškov BI v treh letih



Uporabljena programska oprema



Slika prikazuje nabor Microsoftovih orodij za podporo BI

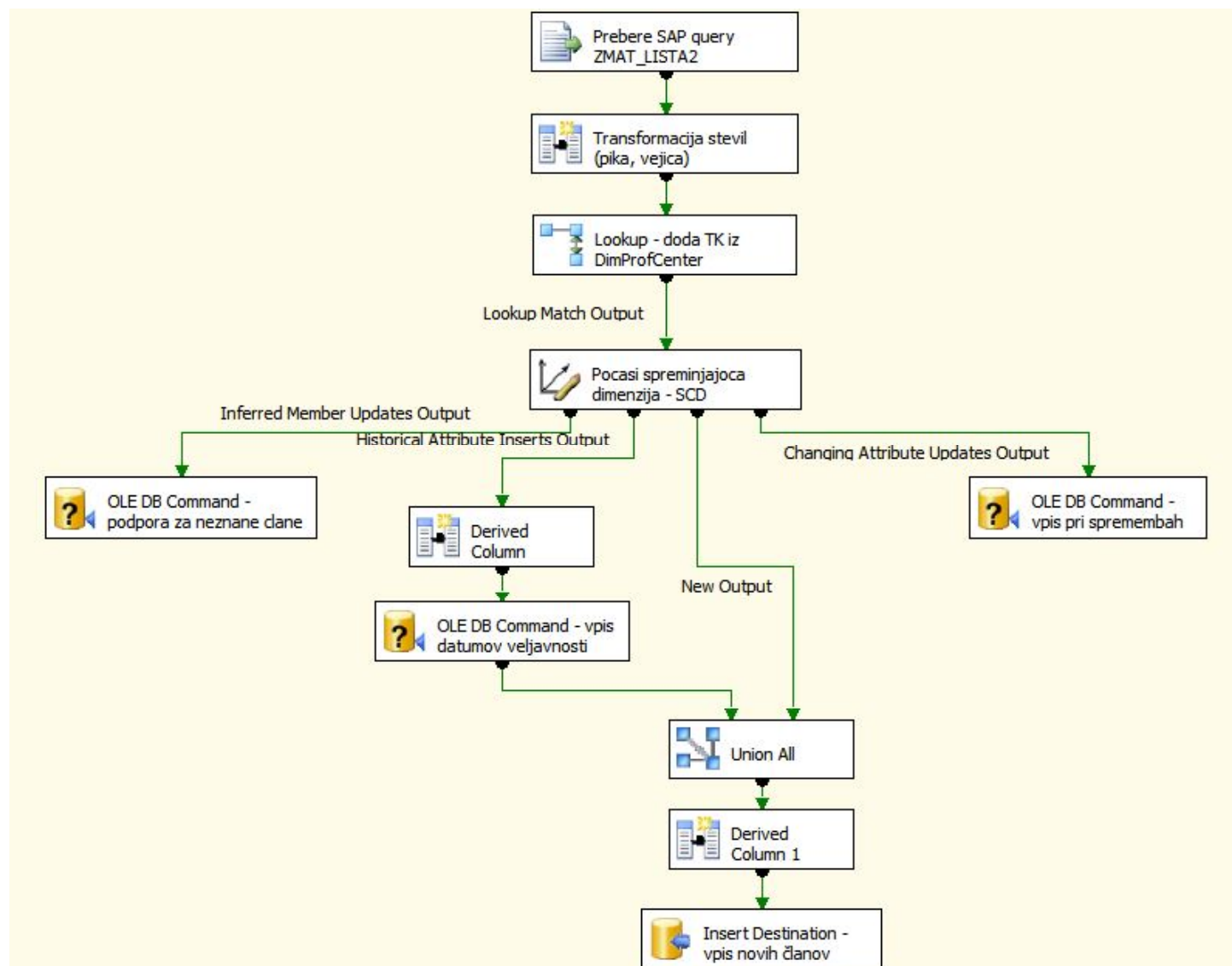
Pri BI gre vedno za set namenskih orodij.

V primeru prototipa so bila uporabljena samo obkrožena.



SSIS – ETL postopek

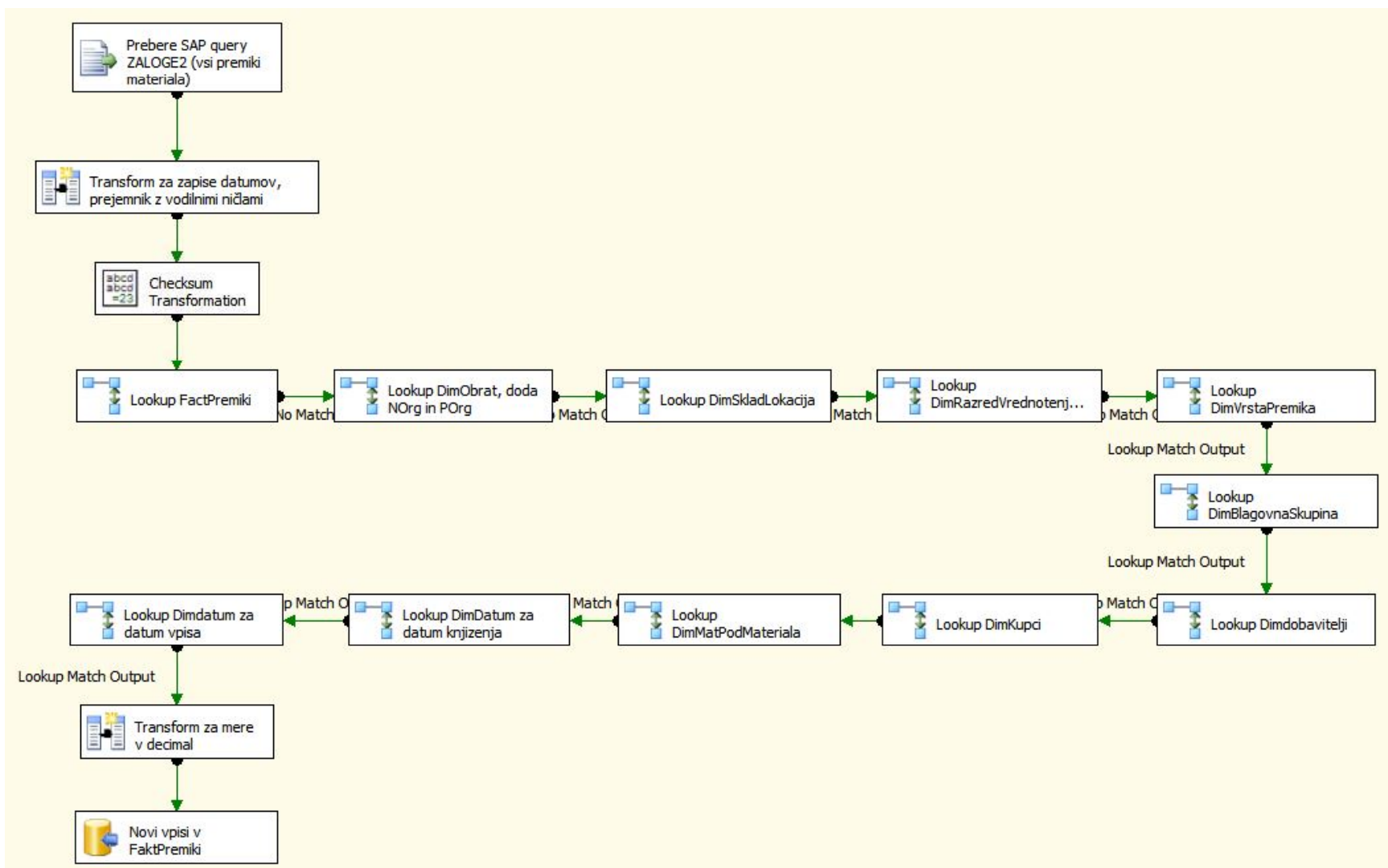
Nalaganje
ene od
dimenzij.





SSIS – ETL postopek

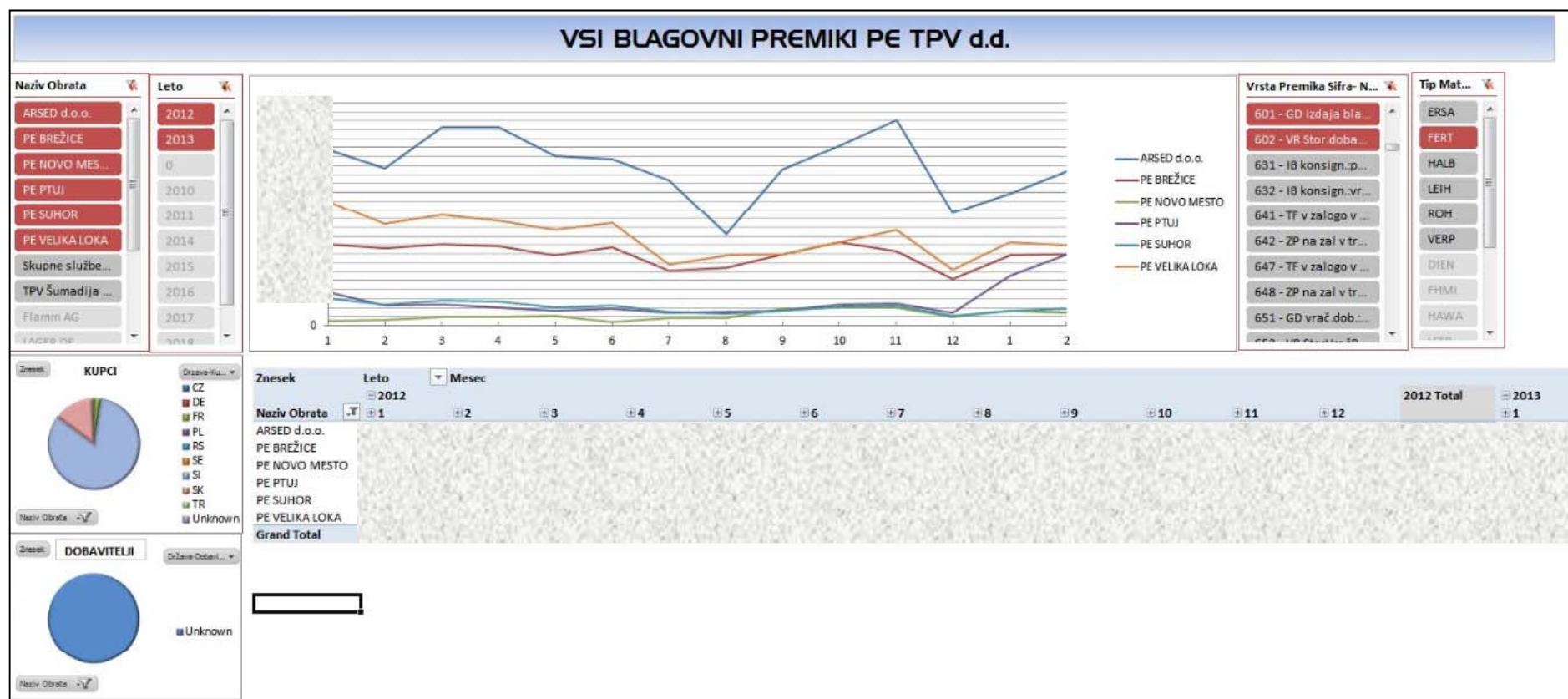
Nalaganje tabele dejstev (*angl. fact table*)





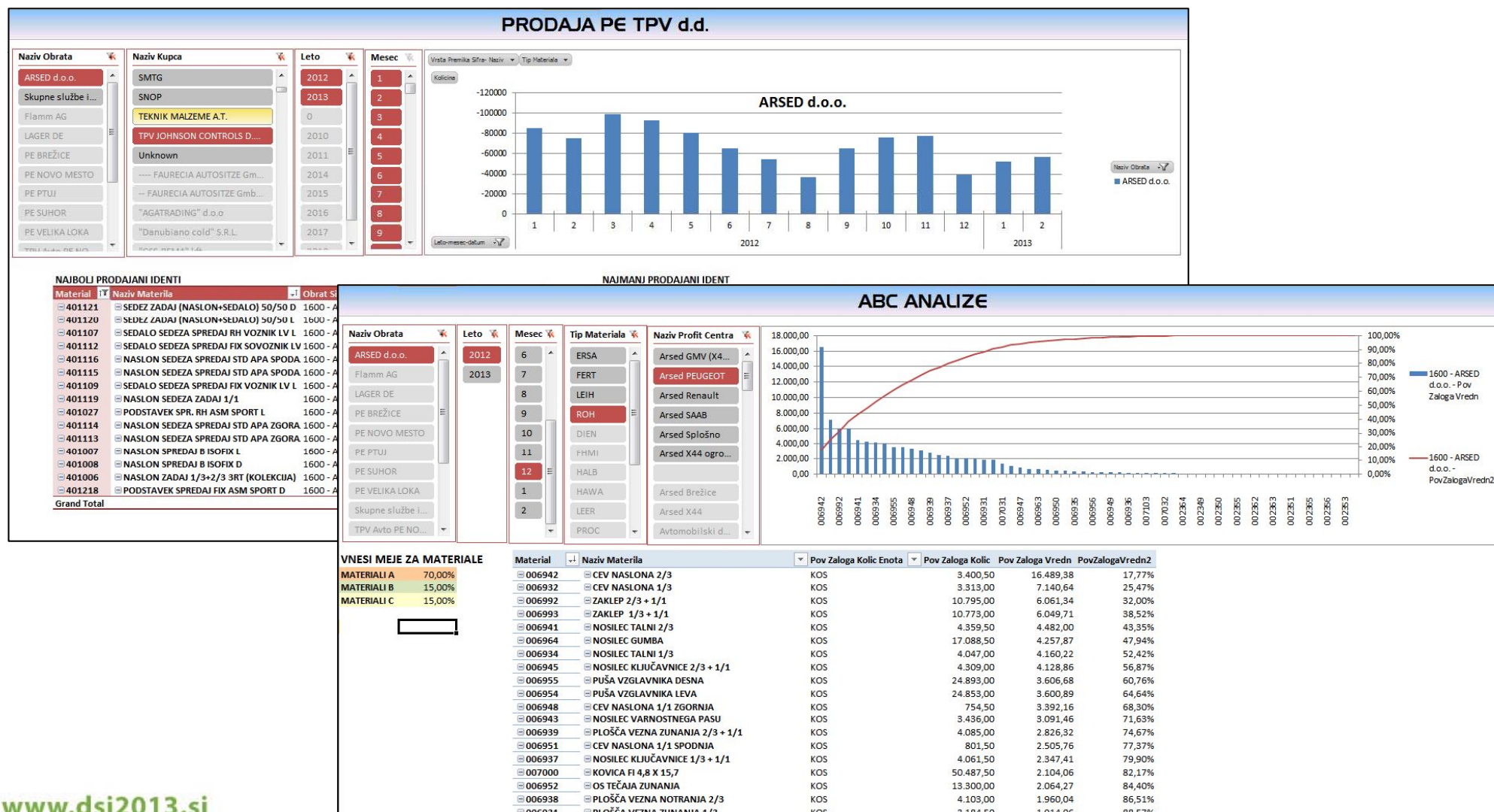
Nadzorne plošče (angl. dashboards, odjemalec Excel)

»Nadzorna plošča je prikazovalnik najpomembnejših podatkov, potrebnih za doseg enega ali več ciljev; urejenih na posameznem zaslonu tako, da so informacije nadzorovane v trenutku« (Few, 2004)





Dobro znano managersko načelo je, da ne moremo upravljati, česar ne moremo izmeriti. Enako resnično je tudi, da ne moremo upravljati, česar ne nadzorujemo. (Shadan, 2005)





Zaključek

Poslovna vrednost poslovne inteligence lahko tudi nič, če z informacijami nič ne storimo. Če jih uporabimo pravilno, pa lahko odločilno poveča poslovno vrednost naših aktivnosti. To se še posebej izrazito pokaže v današnjem hitro se spreminjajočem poslovnem okolju, ki zahteva hitro ukrepanje in veliko prilagodljivost.

Organizacije z vpeljanimi sistemi poslovne inteligence imajo več možnosti za preživetje. Morajo pa te sisteme uporabljati in tudi nadgrajevati. Vodilo in motivator morajo biti vedno poslovni cilji in zahteve poslovnih uporabnikov, ker lahko drugače hitro pridemo tudi do neuspešnih projektov uvedbe



Hvala za vašo pozornost !

Vprašanja?

Pripombe?

Predlogi?