



UNIVERSITETI I EVROPËS LIGLINDORE
УНИВЕРЗИТЕТ НА ЈУГОИСТОЧНА ЕВРОПА
SOUTH EAST EUROPEAN UNIVERSITY

FAKULTETI I SHKENCAVE DHE TEKNOLOGJIVE BASHKËKOHORE
CIKLI I DYTË I STUDIMEVE

IMPLEMENTIMI I DATA WAREHOUSE – RAST STUDIMI NË KOMPANINË KEDS

KANDIDATI:
Bsc. NERTILA KRYEZIU

Tetovë, Maj, 2019

MENTORI:
Assoc. Prof. Dr. FLORIJE ISMAILI

Deklaratë

Deklaroj se kjo tezë dhe puna e raportuar këtu, ka përbërje tërësisht origjinale. Informacioni i marrë nga puna e botuar ose e pabotuar e të tjerëve është e cituar në tekst dhe referencat janë dhënë në listën e burimeve.

Tetovë,
Maj/2019

Falënderim

Pjesa praktike e këtij punimi të magistraturës është zhvilluar në departamentin e IT-së në kompaninë KEDS. I falënderoj drejtorin e departamentit dhe kolegët të cilët më ndihmuan në realizimin e këtij projekti.

Falënderim të veçantë për familjen time të cilët më kanë përkrahur në të gjitha fazat e ngritjes time akademike. Ky rrugëtim nuk do të ishte i mundur pa mbështetjen e tyre.

Faleminderit!

Nertila Kryeziu

Abstrakt

Në botën e biznesit, shumë kompani të mëdha i organizojnë dhe i mirëmbajnë të dhënat e tyre në baza të mëdha të të dhënave. Me kalimin e viteve sasi e madhe e të dhënave do të grumbullohen në sistemin e kompanisë, me rritjen e shpejtë të sistemit ndodhin edhe problemet për krahasimin dhe analizimin e të dhënave. Kompanitë duhet të analizojnë saktë të dhënat e grumbulluara dhe në bazë të tyre të kryejnë disa funksione thelbësore për kompaninë të tilla si planifikimi, parashikimi dhe menaxhimi. Vështirësitë e përmendura më lart kanë detyruar ekspertët e teknologjisë të gjejnë zgjidhje të reja për të ndihmuar njerëzit që marrin pjesë në procesin e vendimmarrjes dhe iu mungojnë njohuritë teknike. Në vitet e fundit, koncepti i data warehouse ka filluar të shfaqet si një lloj i ri i sistemeve komplekse të mbështetjes së vendimeve.

Data warehouse mund të definohet thjeshtë si një grup i të dhënave që mbështetë përdoruesit në marrjen e vendimeve strategjike. Këto të dhëna janë të organizuara në atë mënyrë që do të paraqesin përdorim të lehtë dhe performancë të shkëlqyeshme të raportimit.

Tema e këtij studimi është të hartojë dhe të zhvillojë një data warehouse duke përdorur bazat e të dhënave të kompanisë. Së pari përcaktohen hapat e zhvillimit të data warehouse dhe më pas aplikohet pjesa shpjeguese e funksionimit të saj. Implementimi i data warehouse, që përbën subjektin e kësaj teze, kryhet në një kompani në sektorin e energjisë.

Kjo tezë do të paraqesë përfitimin në kohë dhe efikasitetin e raportimit me përdorimin e data warehouse.

Fjalori

Fjala kyçe	Përshkrimi
EDW	Enterprise Data Warehouse – data warehouse në ndërmarrje
Data warehouse	‘Depo’ e të dhënave, përmbledhë të dhëna nga sisteme të ndryshme.
ETL	Ekstraktim, Transformim dhe Ngarkim i të dhënave nga bazat e të dhënave transaksionale në data warehouse.
SSIS	SQL Server Integration Services.
SSRS	SQL Server Reporting Services.
SSAS	SQL Server Analysis Services.
KPI	Indikatori kyç i performancës
DSV	Data Source View

Përmbajtja

	DEKLARATË.....	II
	FALËNDERIM.....	III
	ABSTRAKT	IV
	FJALORI.....	V
	LISTA E FIGURAVE.....	VIII
1	HYRJE.....	10
	1.1 NDËRLIKIMET	10
	1.2 SHPJEGIMI	11
	1.3 MOTIVIMI	11
	1.4 EKSPERIENCA.....	12
	1.5 DOKUMENTACIONI.....	13
2	DATA WAREHOUSEING.....	18
	2.1 ÇKA ËSHTË DATA WAREHOUSE?	18
	2.2 QËLLIMET E DATA WAREHOUSE.....	19
	2.3 APLIKIMI I DATA WAREHOUSE NËPËR KOMPANI	20
	2.4 ELEMENTET BAZIKE TË DATA WAREHOUSE.....	21
	2.4.1 Sistemi burimor.....	21
	2.4.2 Hapësira e paraqitjes së të dhënave.....	21
	2.4.3 Serveri prezentues	21
	2.4.4 OLAP (On-Line Analytic Processing – Procesimi analitik).....	21
	2.4.5 Aplikacionet për shfrytëzuesit fundorë	22
	2.4.6 Aplikacionet modeluese	22
	2.5 PËRFITTIMET E APLIKIMIT TË DATA WAREHOUSE.....	22
	2.6 FAKTORËT KRYESORË QË MIRREN PARASYSH GJATË DIZAJNIMIT TË DATA WAREHOUSE	23
3	ARKITEKTURA E SISTEMIT	25
	3.1 MODELI DIMENZIONAL	25
	3.2 SHITRESAT E DATA WAREHOUSE	26
	3.3 METADATA	27
	3.4 TREGU I TË DHËNAVE (DATA MARTS).....	28
	3.4.1 Çka është tregu i të dhënave.....	28
	3.5 SHITIGJET E ZHVILLIMIT TË TREGUT TË TË DHËNAVE.....	29
	3.5.1 Shtegu Lart-Poshtë.....	29
	3.5.2 Shtegu Poshtë-Lart.....	30
	3.5.3 Shtegu i Federuar.....	31
	3.6 DALLIMET NDËRMJET DATA WAREHOUSE DHE TREGUT E TË DHËNAVE	
	32	
	3.7 SKEMA YLL.....	33
	3.8 SKEMA SNOWFLAKE	33
	3.9 MBUSHJA DHE PROCESI ETL.....	34

3.10	ALTERNATIVA ETL.....	37
3.11	BIZNIS INTELIGJENCA	38
3.12	RAPORTIMI.....	38
3.13	IMPLEMENTIMI.....	41
3.14	PROTOTIPI I DATA WAREHOUSE SË NDËRMARRJES	41
3.15	KONSTRUKTIMI I ETL.....	42
3.16	MBUSHJA E TABELAVE	43
3.16.1	<i>Mbushja e tabelave dimenzion</i>	43
3.16.2	<i>Mbushja e tabelave fakt</i>	48
3.17	KONSTRUKTIMI I KUBEVE	53
3.17.1	<i>Krijimi i DSV</i>	53
3.17.2	<i>Krijimi i kubit</i>	53
3.17.3	<i>Definimi i kalkulimeve</i>	54
3.17.4	<i>Definimi i KPI</i>	55
3.17.5	<i>Definimi i aksioneve</i>	55
3.18	RAPORTET.....	56
3.18.1	<i>Faturimi i kWh dhe numri i konsumatorëve</i>	56
3.18.2	<i>Faturimi TS</i>	58
3.18.3	<i>Pagesat</i>	59
3.18.4	<i>Raporti i faturave zero</i>	60
4	VLERËSIMI I PERFORMANCËS	61
4.1	PARAQITJA E PËRGJITHSHME.....	61
4.2	PLANI TESTUES.....	61
4.2.1	<i>Faturimi i kWh dhe numri i konsumatorëve</i>	62
4.2.2	<i>Faturimi i TS</i>	63
4.2.3	<i>Pagesat</i>	65
4.2.4	<i>Raporti i faturave zero</i>	66
5	KONKLZIONI	69
	BIBLIOGRAFIA	71

Lista e Figurave

Figura 1. Data warehouse në ndërmarrje	14
Figura 2. Elementet bazë të data warehouse	28
Figura 3. Arkitektura informative e ndërmarrjes.....	29
Figura 4. Shtegu lart - poshtë.....	30
Figura 5. Shtegu poshtë – lart.....	31
Figura 6. Shtegu i federuar	32
Figura 7. Diagrami i bazës së të dhënave të data warehouse.....	43
Figura 8. Control Flow i tabelës dimension 'AccountCode'	44
Figura 9. Data Flow i tabelës dimension 'AccountCode'.....	44
Figura 10. Control Flow i tabelës dimension 'ElDebitors'	45
Figura 11. Data Flow i tabelës dimension 'ElDebitors'	46
Figura 12. Control Flow i tabelës dimension 'PostGLArchive'.....	47
Figura 13. Data Flow i tabelës dimension 'PostGLArchive'	47
Figura 14. Control Flow i tabelës dimension 'Readers'.....	48
Figura 15. Data Flow i tabelës dimension 'Readers'.....	48
Figura 16. Control Flow i tabelës fakt 'Invoices'.....	49
Figura 17. Data Flow i tabelës fakt 'Invoices'.....	49
Figura 18. Control Flow i tabelës fakt 'HPIInvoices'.....	50
Figura 19. Data Flow i tabelës fakt 'HPIInvoices'.....	50
Figura 20. Control Flow i tabelës fakt 'GLArchive'.....	51
Figura 21. Data Flow i tabelës fakt 'GLArchive'.....	51
Figura 22. Control Flow i tabelës fakt 'DOCorrections'.....	52
Figura 23. Data Flow i tabelës fakt 'DOCorrections'.....	52
Figura 24. Krijimi i DSV.....	53
Figura 25. Krijimi i kubit	54
Figura 26. Lidhja e dimensioneve me grupet matëse	54
Figura 27. Kalkulimi	55
Figura 28. Krijimi i KPI.....	55
Figura 29. Krijimi i aksionit ‘drillthrough’.....	56
Figura 30. Faturimi kWh në SSRS.....	58
Figura 31. Raporti i faturimit TS në SSRS	59
Figura 32. Raporti i pagesave në SSRS	60
Figura 33. Raporti i faturave zero në SSRS.....	60
Figura 34. Faturimi kWh në bazën e të dhënave transaksionale.....	62
Figura 35. Faturimi kWh në datawarehouse	63

Figura 36. Rezultatet grafike të faturimit kWh	63
Figura 37. Faturimi TS në bazën e të dhënave transaksionale	64
Figura 38. Faturimi i TS në data warehouse	64
Figura 39. Diagrami i krahasimit të faturimit të TS	65
Figura 40. Pagesat në bazën e të dhënave transaksionale	65
Figura 41. Pagesat në data warehouse	66
Figura 42. Diagrami i rezultateve të pagesave	66
Figura 43. Raporti i zerove në bazën e të dhënave transaksionale	67
Figura 44. Raporti i zerove në data warehouse	67
Figura 45. Diagrami i rezultateve të zerove	68

Kapitulli I

1 Hyrje

1.1 Ndërlikimet

Implementimi i Data Warehouse në kompani, në rastin tonë në kompaninë KEDS, duket se do të zgjidhë problemin e integritit të shumë sistemeve në një burim të përbashkët të të dhënave. Llojet e shënimeve që ka kompania, të dhënat për konsumatorët, punëtorët, faturat, etj., janë të regjistruara në baza të ndryshme të të dhënave. Duke i pasur parasysh këto ndarje, ka të dhëna të cilat janë në interes të më shumë departamenteve mirëpo në to ka qasje vetëm një departament. Të paturit e një data warehouse që i ndanë këto shënime ndërmjet departamenteve do të mund të shkaktoj shkelje të sigurisë. Prandaj, hartimi i një plani për kufizim të qasjeve në të dhëna, është më se i arsyeshëm.

Problemi kryesor i mos pasjes së data warehouse në kompani është konsumimi i të dhënave nga shfrytëzuesit prej formës së raporteve aktuale. Shumë shpesh ata që kanë nevojë për informacion ose kërkojnë informata nga të dhënat që ata i shfrytëzojnë, duhet të presin raportin bazuar në kohën e dikujt tjetër. Atyre iu duhet të manipulojnë me të dhënat të cilat i marrin dhe mundësia e gabimit është më e madhe.

Zhvillimi i teknologjisë kohët e fundit ndihmon shumë në zvogëlimin e gabimeve dhe kompanitë e mëdha duhet të kenë raporte sa më të sakta. Në vend se drejtuesit e kompanisë apo menaxherët të presin marrjen e raporteve nga departamentet përkatëse, atyre duhet t'i u mundësohet të nxjerrin informatat që iu nevojiten në çfarë do kohe dhe pa u varur nga departamentet tjera.

1.2 Shpjegimi

Përdorimi i informatave që ofron data warehouse për marrjen e vendimeve të ndryshme janë komponentë e rëndësishme e biznes inteligjencës. Data warehouse përdoret jo vetëm për të marrë të dhëna, por edhe për t'i paraqitur ato tek shfrytëzuesit që të mund t'iu qasen në formatin e tyre të dëshiruar.

Për shfrytëzuesit të cilët dëshirojnë të marrin të dhënat vetë, pa pritur shpërndarje nga të tjerët ka shumë vegla për raportim të cilat janë në dispozicion dhe të cilat bëjnë që data warehouse të jetë më lehtë të përdorshme.

1.3 Motivimi

Në implementimin e data warehouse në kompani kontribuojnë shumë faktorë. Edhe pse zbatimi i një projekti të tillë mund të kërkojë shumë kohë, burime dhe investime monetare nga ana e kompanisë, ka shumë faktorë motivues për të ecur përpara në zbatimin e këtij projekti.

Motivimi më i rëndësishëm në implementimin e data warehouse është që të kemi një platformë më të përshtatshme për raportim. Duke i kombinuar të gjitha bazat e të dhënave të kompanisë, data warehouse do të jetë burim i vetëm për marrjen e të dhënave nga shfrytëzuesit.

Me përdorimin e data warehouse, përpiluesi i raporteve nuk ka nevojë t'i mësojë të gjitha bazat e të dhënave apo të mundohet t'i lidhë të dhënat ndërmjet tyre, pasi që kjo është një punë shumë e vështirë.

Përveç që kemi një bazë të vetme të të dhënave, kemi edhe shumë përparësi tjera në mos raportimin nga bazat e të dhënave transaksionale. Raportimi nga baza e të dhënave transaksionale mund të ngadalësojë performancën e bazës së të dhënave. Gjatë përpilimit të raportit të dhënat mund të jenë në proces të modifikimit apo fshirjes, të cilat do të bënin raportin jo korrekt.

Një përparësi tjetër e kompletimit të raporteve nga data warehouse është fakti se mund të ketë një vegël të vetme për raportim dhe shpërndarje së raporteve për të gjithë shfrytëzuesit. Përparësia e kësaj lloji vegle do të ishte siguri i shpërndarjes në kohë, menaxhimi i të dhënave si dhe mundësia e ruajtjes së gjurmëve të auditimit të të dhënave.

Nga pikëpamja e departamentit të Teknologjisë Informative, data warehouse do të lehtësojë punën e kontrollit të sigurisë. Në vend se të kontrollojmë sigurinë

në çdo bazë të të dhënave dhe për çdo shfrytëzues, në data warehouse do të kemi vetëm një bazë të të dhënave ku do të menaxhojmë shfrytëzuesit dhe të drejtat e tyre. Për më tepër, IT do të ketë nën kontrollë dhe menaxhim të dhënat e kompanisë të cilat shfrytëzuesit pastaj mund t'iu qasen për të përmbushur nevojat e tyre.

Ndoshta, një prej arsyeve më të nënvlerësuara të data warehouse është fakti se ajo mund të përmbajë të dhëna historike. Një bazë e të dhënave transaksionale nuk përmbanë të dhëna historike sikurse data warehouse sepse shfrytëzuesi shumë herë modifikon apo fshinë këto të dhëna përgjatë ditës.

Arsyeja kryesore për zhvillimin e data warehouse për kompani (Enterprise Data Warehouse) është se ekziston një model i arsyeshëm i të dhënave i zhvilluar për shfrytëzuesit ashtu që qëndrueshmëria është çelësi kryesor për konsumimin e të dhënave në të gjithë kompaninë. Pavarësisht nga burimi i të dhënave, modeli i të dhënave i jep shfrytëzuesit një mënyrë të përbashkët për të marrë të dhënat që ata dëshirojnë për të përmbushur qëllimet e tyre.

Tani më shumë se kurrë më parë, kompanitë mbështeten në informata të sakta, të sigurta dhe të përditësuara në raportet e tyre për të ndihmuar operacionet e kompanisë së tyre. Duke pasur një burim të përbashkët të të dhënave për shfrytëzuesit dhe duke larguar barrën e raportimit direkt nga bazat e të dhënave transaksionale mund të përmirësohet efikasiteti dhe ndarja e të dhënave në të gjithë kompaninë.

Zgjidhja e data warehouse zëvendëson Excel-in dhe platformat tjera raportuese me zgjidhje moderne dhe të centralizuara të raportimit dhe analizës. Aftësia për të konsoliduar të gjitha të dhënat e rëndësishme dhe të lidhura për një verzion të vetëm të së vërtetës është themelore për raportimin e saktësisë.

1.4 Eksperienca

Unë do të paraqes një shembull real të bartjes së raportimit prej bazës së të dhënave transaksionale në data warehouse. Ky shembull krahason diferencën ndërmjet ekzekutimit të pyetësorit të njejtë në tabela në bazën e të dhënave transaksionale kundrejt pyetësorit me të dhëna të grupuara në skemën yll që përmbajnë të dhënat e njejta.

Duke përodrur një sërë masash në kohën e duhur, avantazhet e raportimit nga data warehouse do të jenë më të qarta. Krahasimet do të tregojnë për dallimet

në kohë për pyetësorë të ndryshëm, të cilët do të kthejnë të përmirësuar kohën e përgjithshme të raportimit nga shfrytëzuesit. Ka avantazhe dhe disavantazhe për secilin, por në përgjithësi, do të jetë shumë e dobishme raportimi nga data warehouse. Një nga përparësitë kryesore të implementimit të EDW është për të përmirësuar kohën e raportimit dhe të prezantohen të dhënat në mënyrë që shfrytëzuesit mund t'i shfrytëzojnë ato për të bërë vlerësime më të mira dhe të përmirësojnë efikasitetin brenda departamentit të tyre.

1.5 Dokumentacioni

Qëllimi i këtij punimi është të tregohen përparësitë e implementimit të EDW në kompani. Me EDW ne kemi mundësi t'i integrojmë të dhënat nga shumë baza të të dhënave të cilat shfrytëzohen nga aplikacione të ndryshme nga departamentet e kompanisë. Me gjithë këto informata të përhapura në të gjithë kompaninë, duke përfshirë edhe mundësinë që një person mund të ketë qasje në më shumë se një sistem, të kemi vetëm një burim të të dhënave është një arsye e rëndësishme për implementimin e EDW. Departamentet raportojnë vetëm brenda njësie të tyre, kështu që ka mundësi që mos të bëhet shkëmbimi i informacioneve në mënyrë adekuate nëpër kompani. Edhe në qoftë se informacioni tashmë është shpërndarë, shumë përdorues do të shkëmbejnë materiale përmes e-mailit apo ndonjë forme tjetër që nuk është edhe aq e përshtatshme dhe mund të dështojnë në ndarjen e materialit në mënyrë të arsyeshme dhe të sigurtë. Duke i pasur të dhënat në një burim të vetëm dhe të menaxhuara mirë nga departamenti i IT-së, të dhënat janë në dispozicion për shfrytëzuesit në një platformë të vetme dhe të gatshme për ata që kanë nevojë t'i marrin këto informata.

Kompania ka shumë departamente dhe në secilin prej tyre kemi aplikacione të cilat kanë baza të të dhënave që i ruajnë të gjitha informatat e përdorura nga aplikacioni. Bazat e të dhënave dhe aplikacionet mund të jenë të mëdha apo të vogla, mirëpo secila prej tyre do të ketë vend në EDW për të përmbushur nevojat e shërbimeve të raportimit brenda kompanisë.

Gjatë shqyrtimit të nevojave për EDW është e rëndësishme të shikohen të gjitha departamentet dhe të shikohet se çka kanë mundësi ata të ndihmojnë në këtë projekt.

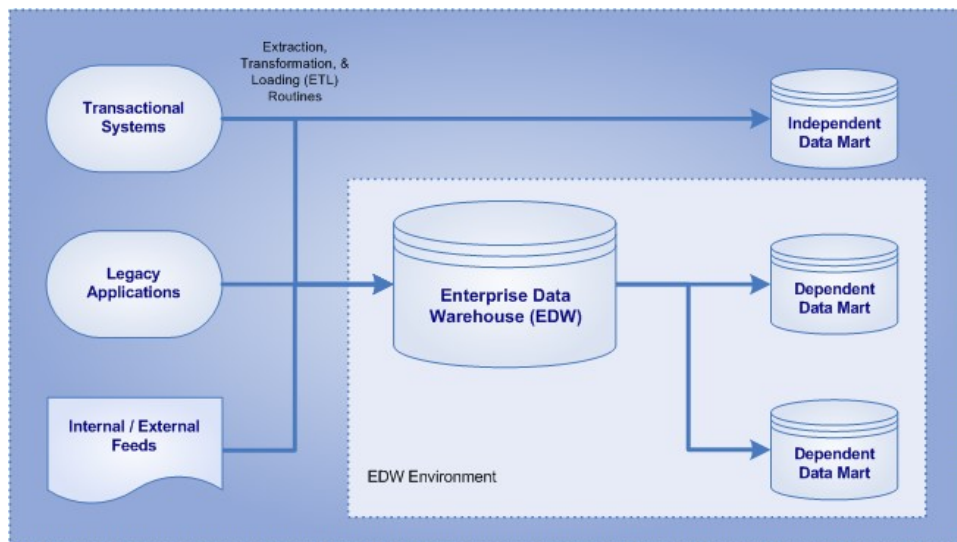


Figura 1. Data warehouse në ndërmarrje

Me krijimin e EDW mund të lindin pyetje të ndryshme se çfarë përfitimi do të kemi nëse i kemi të dhënat të mbledhura në një vend, veçanërisht nëse kompania funksionon mirë edhe në atë mënyrë. Ideja e EDW do të jetë ofrimi i raportimit më të lehtë dhe qasja më e shpejtë në informata.

Duke i pasur të dhënat nga baza të të dhënave të ndryshme në një raport, çon në përhapje më të mirë të informatës dhe zvogëlon mundësinë e implikimeve të performancës.

Për implementimin e EDW nuk ka justifikim më të madh se mundësia e konsumimit dhe raportimit të të dhënave në mënyrë më të mirë dhe më të shpejtë se forma aktuale në të cilën janë mësuar të punojnë përdoruesit. Me veglat e raportimit të cilat janë në dispozicion tani, mund të mendohet se gjërat nuk mund të jenë më mirë se që janë tani. Megjithatë, siç e kemi shqyrtuar implementimin e data warehouse, do të gjejmë se ka zgjidhje më të mira për raportim. Më tej, siç i kemi analizuar ne, do të ketë përmirësim të dukshëm të infrastrukturës së data warehouse.

Me mundësitë e mëdha që kompanisë i ofron EDW, është e rëndësishme të kujtojmë se për të ndërmarrë një përpjekje të tillë kërkohet një grup i caktuar njerëzish brenda kompanisë që nevojiten për krijimin e sistemit. Këtu përfshihet një përkrahës i fuqishëm ekzekutiv që do të jetë udhëheqës, një arsye solide e kompanisë që ofron motivimin, dhe në fund një projekt i organizuar i cili duhet

të ketë një lider menaxhmenti. Çdo pjesë e këtij koordinimi për të zhvilluar data warehouse është më sfiduese se zhvillimi fizik i vetë data warehouse, pasi motivimi dhe përkrahja financiare është më se e nevojshme.

Një nga pjesët më të rëndësishme të implementimit të EDW është se duhet të ketë një përkrahës të fuqishëm ekzekutiv që do ta përkrahë përpjekjen për kalim në data warehouse. Përkrahësit ekzekutiv definojnë si: “përkrahës të fuqishëm të biznesit të cilët kanë influencë të madhe në kompani” [1].

Gjithashtu i rëndësishëm është edhe virtyti i udhëheqësit, i cili është jetik për udhëheqjen e të tjerëve drejt hapave të ndryshëm të zhvillimit të data warehouse. Për më tepër, ”përkrahësit efektiv janë në gjendje të mirren me problemet dhe pengesat afat-shkurtëra për shkak se ata janë të fokusuar në suksesin afatgjatë të projektit“ [1]

”Kompanitë të cilat i implementojnë me sukses zgjidhjet DW/BI, në mënyrë tipike kanë një sens të fuqishëm të urgjencës për përmirësimin e qasjes në informata të shkaktuara nga motivet bindëse të biznesit“ [1].

Kompanitë mund të jenë të suksesshme me metodat e tyre aktuale për nxjerrjen e informatave, mirëpo ka gjithmonë vend për përmirësim dhe dëshirë për kursim në kohë dhe rezultate më të shpejta të cilat i ofron implementimi i data warehouse.

“Sistemet DW/BI që lidhen me motivimet dhe iniciativat strategjike të biznesit janë një shansë e mirë e suksesit“. [1] Kjo është një arsye për të besuar se nëse një proces i përmirësuar i raportimit do t’i plotësojë më mirë nevojat e biznesit, atëherë implementimi i data warehouse do të bëhet prioritet.

Përderisa kemi mbështetjen e përkrahësve të kompanisë për të krijuar EDW, atëherë i mbetet departamentit të IT-së të ofroj zgjidhje për implementimin e data warehouse. IT-së i mbetet të trajtojë stafin dhe të mbikëqyr projektin deri në kompletimin e saj. Drejtuesit të projektit i duhet të udhëheqë stafin e dy grupeve, atë të IT-së dhe departamenteve tjera.

Ekipi i projektit do të përbëhet nga ata që kuptojnë se si implementohet një projekt i tillë apo do të kenë një bazë kontraktuese për të ndihmuar në zhvillimin e data warehouse. Fakti se “Është e zakonshme për ekipin e DW/BI të merr pjesë në detyrat e zhvillimit dhe operimit me të dhëna duke pasur parasysh natyrën përsëritëse të ciklit të zhvillimit të projektit DW/BI“ [1], tregon se ekipi i

projektit është i përbërë më shumë nga staf teknik por edhe përkrahës të cilët do t'i ndihmonin në edukimin e përdoruesve. Ata do t'i ndajnë njohuritë e tyre përmes diskutimeve të hapura apo mbledhjeve të organizuara.

Pas takimit të parë rreth projektit të përbërë nga të gjithë përdoruesit dhe udhëheqësit, ekzekutivi i projektit fillon me planin e projektit dhe grumbullimin e kërkesave. Kërkesat për data warehouse mund të duken të qarta sipas shpjegimit të deritashëm, mirëpo në këtë fazë ka vend për të dhëna nga të gjithë të përfshirët në projekt. Këto të dhëna mund të grumbullohen përmes diskutimeve në të cilat përmbledhen nevojat e të gjithëve rreth këtij projekti. Ralph Kimball thekson se në këtë fazë nuk është e nevojshme vetëm të dihet se çka dëshirojnë të tjerët nga data warehouse, por më shumë të mësohet se si lidhet ajo me punën e tyre. Ai thotë se "Ju duhet të flisni me ta rreth punës, objektivave dhe sfidave të tyre" [1]. Kjo tashmë ka kuptim pasi data warehouse do të përfitojë më mirë komunitetin nëse ua bënë punën më të lehtë dhe u ofron zgjidhje për disa nga punët e tyre që janë më të vështira. Pas përfundimit të kësaj serie të grumbullimit të informatave, plani i projektit vazhdon me vendosjen e qëllimeve se si do të përfundojë projekti nëpër fazat e nevojshme për të qenë i suksesshëm. "Koordinimi i përkrahjes, administrimit dhe praktikave më të mira në nivel të programit janë vendimtare për të rritur kthimin potencial të biznesit nga investimi i juaj në DW/BI" [1].

Për bartjen e të dhënave nga një burim në tjetrin nevojitet një hap tjetër i rëndësishëm para se të bëhet kjo bartje, dhe kjo duhet të jetë si projekt i veçantë në nivel departamenti, e jo si pjesë e implementimit të data warehouse.

Ekzistojnë shumë hapa dhe procedura të nevojshme për t'u ndjekur gjatë bartjes së të dhënave në data warehouse, të cilat zakonisht janë më të lehta të ndiqen në fazat e hershme të zhvillimit.

Gjatë bartjes së të dhënave duhet të ekzistoj një standard i cili do të sigurojë të dhëna më të sigurta. Duke pasur një plan të standardizimit para futjes së të dhënave në vend se t'i pastrojmë ato më vonë, do të jetë një zgjidhje më e mirë, në këtë rast do të bëhet ndryshimi i sjelljes së njerëzve në vend të sistemit. Standardizimi do të nevojitet edhe gjatë fazës së zhvillimit të ETL.

Para se të fillohet procesi i konstruktimit të EDW, kemi shumë pjesë të rëndësishme të projektit që paraqesin sfondin e data warehouse dhe do të çojnë në implementimin e suksesshëm të këtij projekti.

Pjesa më e rëndësishme e projektit është rezultati përfundimtar i data warehouse ku shihet se si e gjithë përmbajtja e departamenteve vjen e bashkuar për t'iu dhënë mundësinë përdoruesve të raportojnë një pasqyrë të madhe.

Para se të filloj përdorimi i data warehouse, është e nevojshme që lideri të përkrahë data warehouse dhe të përhapë vizionin e projektit te ekipet e përdoruesve që do të përdorin data warehouse. Gjithashtu shumë e rëndësishme është të kemi një zyrë të menaxhimit të projektit ku do të kryhet ekzekutimi i projektit. Përdoruesit e sistemit apo pjesëtarët e ekipit do të jenë në gjendje të japin mendimet e tyre se si ata do të zgjidhnin sfidat e tyre aktuale. Ata gjithashtu do të marrin si detyrë pastrimin apo standardizimin e të dhënave aktuale të tyre.

Rezultatet e të pasurit data warehouse duhet të justifikojnë investimet kohore dhe shpenzimet monetare, dhe një kuptim shumë i qartë i sfondit është i nevojshëm për të kuptuar të gjitha implikimet e projektit para implementimit.

Kapitulli II

2 Data warehouseing

2.1 Çka është data warehouse?

Data warehouse është një depo e të dhënave të bashkuara, për t'i mundësuar përdoruesit fundor analizë të shpejtë të të dhënave. Këto të dhëna mirren nga burime të ndryshme të cilat ruhen në mjedise të ndryshme fizike. Data warehouse është një mjedis i cili përfshinë të gjitha aplikacionet dhe operacionet për të menaxhuar procesin e grumbullimit të të dhënave dhe dërgimin e tyre tek përdoruesit, e jo vetëm një bazë e të dhënave relacionale që përmbanë të dhëna historike të derivuara nga të dhënat transaksionale.

Data warehouse nuk ka një definicion standard dhe njerëzit që punojnë në të e kanë definuar në disa mënyra si më poshtë:

”Arkitektura bazë e data warehouse paraqet një depo të cilën ne e mendojmë si një sistem i madh dhe i vetëm i ruajtjes, i cili vlerësohet përafërsisht si një ndërmarrje e të dhënave, që ndërmjetëson përdoruesin fundor dhe burimet e të dhënave prodhuese“ [2].

“Data warehouse është kopje e të dhënave transaksionale të strukturuar në mënyrë specifike për raportim“ [1].

“Data warehouse është grumbullim i të dhënave i integruar, 'subject-oriented', i ndryshuar në kohë, dhe i paqëndrueshëm në mbështetje të procesit të vendimmarrjes së menaxhmentit“ [3].

Sipas William Inmon:

- **subject - oriented.** Data warehouse është i dizajnuar për të ndihmuar në marrjen e vendimit për një subjekt specifik. Psh., aplikacionet për të dhënat e shitjeve përmbajnë shitje specifike të produkteve specifike për

konsumatorë specifike. Të dhënat e orientuara në subjekt i japin një imazh stabil procesit të biznesit, të pavarur nga sistemi i trashëgimisë. Me fjalë të tjera, kap natyrën bazë të mjedisit biznesor.

- **integruar.** Data warehouse përbëhet nga lloje të ndryshme të të dhënave të cilat grumbullohen nga sisteme të ndara dhe kjo mund të shkaktoj konflikt dhe mospërputhje në njësitë matëse. Për shkak të kësaj ato duhet të vendosen në format konsekuent dhe kështu ato paraqiten të integruara.
- **nonvolatile.** nonvolatile nënkupton që të dhënat të cilat futen një herë në data warehouse ato më nuk mund të ndryshojnë. Kjo është logjike pasi qëllimi i warehouse është që t'i mundësoj përdoruesit të analizoj se çfarë ka ndodhur. Të dhënat e reja gjithmonë bashkangjiten në bazë të të dhënave e cila vazhdimisht absorbon të dhëna të reja, duke i integruar me të dhënat ekzistuese.
- **i ndryshueshëm në kohë.** në pikën e ndryshimit në kohë ka ndryshim ndërmjet të dhënave operacionale dhe informacionale. Të dhënat operacionale janë valide vetëm në momentin e qasjes. Kur kërkohen nevojat e performancës nevojiten të dhënat historike. Të dhënat në data warehouse paraqesin të dhëna për një periudhë të gjatë kohore; analizat historike mund të performohen më lehtë [3].

2.2 Qëllimet e data warehouse

Qëllimet themelore të data warehouse janë:

1. "I bën informatat e organizatës të qasshme." Të dhënat në data warehouse janë të qarta dhe të etiketuara me korrektësi. Mënyra e qasjes në të dhëna është shumë e lehtë dhe shumë e shpejtë. Këto cilësi mund të listohen si: të kuptueshme, të navigueshme, dhe me performancë të lartë.
2. "I bën informatat e biznesit konsekuente." Informatat konsekuente kanë rëndësi kyçe në data warehouse përderisa të dhënat grumbullohen nga pjesë të ndryshme të biznesit. Ato duhet të harmonizohen plotësisht. Nëse dy njësi të biznesit nuk nënkuptojnë të njëjtën gjë atëherë ato etiketohen në mënyrë të ndryshme.
3. "Të jetë burim i informatave të përshtatshme dhe elastike." Kjo mundëson të shtohen të dhëna të reja dhe të bëhen pyetësorë të rinj pa pasur nevojë për ndryshim të të dhënave dhe teknologjive ekzistuese.
4. "Të jetë bastion i sigurtë për ruajtjen e asetëve të pronarit." Data warehouse përveç kontrollimit të qasjes efektive në të dhëna, kompanisë i

jep një dukshmëri të mirë në përdorimin e atyre të dhënave, madje edhe pas largimit të tyre nga data warehouse.

5. “Të jetë bazë për marrjen e vendimeve.” Data warehouse siguron të dhëna të sigurta për marrjen e vendimeve. Vendimet janë paraqitje e data warehouse [1].

2.3 Aplikimi i Data Warehouse nëpër kompani

Më poshtë janë tri shembuj të data warehouse në kompani.

Kompani e shitjes

- Analizë e shitjeve
- Parashikim
- Përcjellje e inventarit

Prodhimtari e teknikës

- Matja dhe analiza e zingjirit të furnizimit
- Përfitimi i konsumatorit
- Partneriteti strategjik – negociimi me furnizuesin dhe konsumatorin
- Negociimi për kontrata
- Përfitimi sipas biznesit
- Analiza specifike për produktivitet dhe kualitet

Kompani për prodhimin e ushqimit

- Matja kundrejt konkurrencës
- Mundësia e projektimit si do të jetë produkti i ri
- Analiza e shitjeve sipas regjionit dhe dyqanit
- Mundësia e paraqitjes se si shkon shitja e produkteve në dyqanet konkurrencte, në dyqanet e të njejtit zingjir dhe kundrejt produkteve konkurrencte

- Analiza e kostos së prodhimit dhe performancës
- Analiza e produktivitetit
- Analiza e niveleve të artikujve
- Planet dhe parashikimet [4]

2.4 Elementet bazike të Data Warehouse

2.4.1 Sistemi burimor

Sistem burimor quhet sistemi trashëgimor i cili tërheqë të dhënat biznesore dhe transaksionet. Sistemi burimor duhet të jetë në dispozicion dhe jep mundësinë për të shpërndarë dimensionet bazike si të produktit dhe të konsumatorit me sistemet tjera trashëguese të kompanisë. Është burimi më i madh i të dhënave për sistemet e analizave prandaj mundëson krijimin e pyetësorëve dhe raporteve të menaxhmentit direkt nga këto sisteme.

2.4.2 Hapësira e paraqitjes së të dhënave

Hapësira e paraqitjes së të dhënave është hapësira fillestare e ruajtjes ku proceset që pastrojnë, transformojnë, kombinojnë, arkivojnë aplikohen në të dhëna në mënyrë që të përdoren pastaj në data warehouse. Hapësira e paraqitjes së të dhënave vepron si një urë ndërmjet sistemit burimor dhe serverit prezentues. Kufizimi kryesor i hapësirës së paraqitjes së të dhënave është se ajo nuk ofron pyetësorë dhe shërbime prezentuese.

2.4.3 Serveri prezentues

Serveri prezentues është një makinë fizike që ruan të dhënat e procesuara për kërkesat e shfrytëzuesve fundorë për pyetësorë dhe raporte. Ai furnizohet nga hapësira e paraqitjes së të dhënave. Nëse burimi i prezentimit të pyetësorëve për të dhënat e kompanisë organizohet rreth modelit me entitet relational, kuptueshmëria dhe performanca do të humbasin. Tabelat do të organizohen si skemë yll nëse serveri prezentues prezenton dhe ruan të dhënat në strukturë dimensionale.

2.4.4 OLAP (On-Line Analytic Processing – Procesimi analitik)

OLAP mundëson krijimin e pyetësorëve dhe prezentimin e të dhënave qofshin ato tekst apo numra, nga data warehouse për shfrytëzuesit fundorë. Teknologjia OLAP bazohet në kubin multidimensional të të dhënave dhe bazat e të dhënave OLAP kanë strukturë multidimensionale.

2.4.5 Aplikacionet për shfrytëzuesit fundorë

Këto aplikacione ndihmojnë shfrytëzuesit fundorë të krijojnë pyetësorë, të bëjnë analiza dhe të përmbushin aktivitete të tjera të cilat i përmbushin nevojat e kompanisë siç janë veglat për qasje në të dhëna dhe veglat për pyetësorë ad hoc.

Veglat për qasje në të dhëna për shfrytëzuesin fundor punojnë me sesion SQL dhe i ofrojnë shfrytëzuesit raport, shfaqje të të dhënave apo formë tjetër të analizës.

Vegla për pyetësorë ad hoc lehtëson përgatitjen e pyetësorëve duke i ofruar shfrytëzuesit mundësinë e përdorimit të pyetësorëve shabllon.

2.4.6 Aplikacionet modeluese

Aplikacionet modeluese mundësojnë transformimin apo përmbledhjen nga data warehouse sipas modelit të parashikimit, modelit të sjelljes së rezultateve, modelit të shpërndarjes dhe veglave për data mining.

2.5 Përfitimet e aplikimit të data warehouse

Përfitimet kryesore nga data warehouse janë listuar më poshtë:

- DW ofron qasje në informatat administrative për marrësit e vendimeve.
- I bën analizat e të dhënave më të lehta dhe më të shpejta. DW rrit produktivitetin e marrësve të vendimeve të kompanisë.
- Ekstraktimi i të dhënave nga burimi i tyre original në hapësira qendrore zgjidh problemin e performancës i cili lind gjatë nxjerrjes së analizave komplekse nga të dhënat operationale.
- Të dhënat në DW janë të ruajtura në mënyrë të specializuar dhe quhet bazë e të dhënave multidimensionalë. Kjo formë i bën më efektive dhe më të shpejtë pyetësorët e të dhënave.
- DW grumbullon sasi shumë të madhe të të dhënave. Në krahasim me bazat e të dhënave relacionale të dhënat në DW nuk është e nevojshme të jenë të normalizuara, në mënyrë që të shfaqen më shpejtë.

2.6 Faktorët kryesorë që mirren parasysh gjatë dizajnit të data warehouse

Faktorët kryesorë të cilët duhet të mirren parasysh gjatë dizajnit të DW janë:

Madhësia dhe arkitektura: kjo i referohet asaj se sa dëshiron kompania të jetë i gjerë DW i tyre dhe tërësisht mvaret nga nevojat e kompanisë. Por zgjidhja dhe plani duhet të bëhen sipas programit implementues në të cilin madhësia e projektit është konform madhësisë dhe nevojave të kompanisë. Duke pasur një plan të pararendinuar për madhësinë dhe arkitekturën e data warehouse ndihmon në mënjanimin e problemit “scope creep” i cili është shkaktari kryesor në dështimin e projektit të data warehouse.

Ndërveprimi: nëse kompania përdorë bazën e të dhënave në SQL, Oracle apo ndonjë tjetër bazë të të dhënave specifike të cilës i duhet të bashkëveprojë me ndonjë tip tjetër të bazës së të dhënave, atëherë çdo DW duhet të ofrojë ndërveprim ndërmjet formateve dhe platformave të bazës së të dhënave. E njëjta vlenë edhe për ndërveprimin ndërmjet sistemeve operative në nivelin e shfrytëzuesve fundorë: nëse jeni duke përdorur sistemin operativ Unix për menaxhim të një baze të të dhënave dhe Windows për një bazë të të dhënave tjetër duhet të vendoset për një zgjidhje e cila siguron se do të ketë ndërveprim të sistemeve operative.

Kostoja: kostoja është ndër faktorët kryesorë gjatë dizajnit të DW dhe pjesa e parë e projektit të DW e cila mund të shkojë keq nëse nuk planifikohet në përputhje me burimet e mundshme reale. Kostoja është e rëndësishme edhe në pikëpamjen strategjike kur bëhet planifikimi i madhësisë së DW. Kostoja i lejon kompanisë të bëjë zgjidhjen më relevante për nevojat e tyre, dhe të bëjë shfrytëzimin më të mirë të fondit duke evituar funksionalitetet e panevojshme nga DW.

Menaxhimi: kompania duhet të vendosë saktësisht se si duhet të ndahen përgjegjësitë për projektin e DW, dhe cilat njësi administrative do të jenë përgjegjëse për cilin lloj të informatave. Këto kompetenca duhet të qartësohen para zhvillimit ashtu që të mos ketë konfuzione se kush për çfarë është përgjegjës.

Mbështetja: nevoja për trajnime efektive gjithëpërfshirëse shpesh injorohet dhe mund të kushtojë deri në pesëmbëshjetë për qind të buxhetit të zhvillimit,

varësisht nga projekti në fjalë. Trajnimi është ndër gjërat e para të cilat duhet të mirren parasysh gjatë planifikimit të buxhetit për DW dhe është vendimtar në suksesin e projektit. Rekomandohet që të jetë një koordinator i trajnimeve që do të ketë në menaxhim gjithë procesin duke përfshi kohën, lokacionin, aranzhimin e materialeve, etj.

Mbështetja institucionale: profesionistët e teknologjisë përveç zgjedhjes së administratorëve dhe shfrytëzuesve duhet të sigurojnë edhe mbështetje të vizionit dhe procesit. Suksesi përfundimtar i DW mvaret shumë nga mbështetja e shfrytëzuesve fundorë brenda njësive të kompanisë.

Interpretimi dhe Funksionaliteti: siç është cekur nga Margaret Lester “Faktori i cilësisë së interpretimit në DW është në atë se sa do të kuptojnë shfrytëzuesit informatat nga raportet e DW“. Nëse administratorët nuk mund të kuptojnë përfitimin e shfrytëzuesve nga informatat dhe nëse informata nuk është interpretuese apo funksionale për shfrytëzuesit, kjo shkakton dyshim në vlerën e të dhënave të paraqitura.

Shtimi dhe fleksibiliteti: edhe pse shumë data warehouse paraqiten me një numër modest të qëllimeve fillestare, dizajni fillestar duhet të llogaritet mënyrën se si do të ndryshojnë këto qëllime kur ndryshojnë qëllimet e kompanisë. Çdo arkitekturë e data warehouse duhet të jetë në gjendje të ndryshohet pa shumë vështirësi dhe pa investime shtesë në kohë dhe kosto.

Privatësia/Siguria: shkalla në të cilën privatësia dhe siguria do të jenë çështje më të vogla ose më të mëdha mvaret nga definimi i shtrirjes. Të gjitha informatat duhet të jenë të mbrojtura në mënyrën më të mirë të mundshme dhe tipet e sigurisë dhe llojet e problemeve të privatësisë ndryshojnë varësisht se çfarë përmbanë informata në një data warehouse të mundshme. Arkitektura e paradefinuar e sigurisë gjithashtu u mundëson menaxherëve kontrollë më të rreptë mbi të dhënat, dhe u lejon atyre të sigurojnë më mirë biznesin. [5]

Kapitulli III

3 Arkitektura e sistemit

3.1 Modeli dimenzional

Modeli dimenzional i cili është i dizajnuar për të rritur performancën e pyetësorit dhe të jetë më i kuptueshëm, është një model alternativ i modelit me entitet relacional. Modeli dimenzional përbëhet nga tabela fakt dhe tabelat dimension.

Tabela fakt përmbanë matjet e kompanisë që preferohet të jenë numerike dhe shtuese. Aty duhet të ketë një set prej dy e më shumë çelësave të jashtëm që ndihmojnë lidhjen ndërmjet tabelës fakt dhe atyre dimension.

Tabela dimension është tabelë plotësuese e tabelës fakt. Shumë prej tyre janë me tipare tekstuale. Ajo gjithashtu ka çelësin kyç që mundëson lidhjen me tabelën fakt.

Procesi kryesor i organizimit dhe konfigurimit të të dhënave është faza e njohur si ‘modelimi dimenzional’, që është hapi që organizon dhe konfiguron destinacionin final të të dhënave. Kimball thekson: “Modelimi dimenzional është pranuar gjerësisht si qasje e preferuar për prezantim të DW/BI. Praktikuesit dhe ekspertët kanë pranuar se të dhënat e prezentuara në veglat e BI duhet të bazohen në thjeshtësi për të pasur shanse për sukses“ [1]. Kjo pohon se ai është më logjiku dhe më i përhapuri në praktikë për të transformuar të dhënat duke i denormalizuar ato në modelin dimenzional të përbërë nga tabelat fakt dhe dimension. Edhe pse është e mundur që me përdorimin e ETL të barten të dhënat në data warehouse dhe të bëhen të prezentueshme në të njëjtën formë si burimi, kjo nuk ka shumë kuptim dhe dështon në marrjen e avantazheve të benefiteve të arkitekturës së pranuar të data warehouse.

Për më tepër, nëse ndokush bënë përpjekje për të ndërtuar një data warehouse, ai arsyeton se shfrytëzuesit fundorë do t'i kenë në dispozicion të dhënat e tyre në mënyrën më të thjeshtë për BI të tyre.

3.2 Shtresat e data warehouse

Brenda fushëveprimit të data warehouse “ka katër komponente të ndara që mirren në konsideratë gjatë shqyrtimit të data warehouse – sistemi operativ i burimit, fusha e përqendrimit të të dhënave, fusha e prezentimit të të dhënave dhe veglat për qasje në të dhëna” [1].

Burim operativ janë bazat e të dhënave transaksionale në prodhim që kanë qenë lokacioni për të gjitha të dhënat tona.

Fusha e përqendrimit të të dhënave është një bazë e të dhënave e kufizuar për shfrytëzuesit, mirëpo është ndalesa e parë e të dhënave nga sistemi transaksional prej ku të dhënat vijnë së bashku nëpërmjet ETL dhe bëjnë transformimin e tyre para se të vazhdojnë në data warehouse aktual.

Fusha e prezentimit të të dhënave është data warehouse aktual i mbushur me të dhëna që janë në tabelat fakt dhe dimenzion, metoda standarde e formatimit të data warehouse. Shtresa e prezentimit i paraqet të dhënat shfrytëzuesve në formatin e dëshiruar për të marrë vendime me efektivitet.

Arkitektura fizike e data warehouse fillon në shtresën e përqendrimit të të dhënave të modelit të data warehouse. Vegla ETL e cila mundëson ekstraktim, transformim dhe ngarkim të të dhënave nga burimet e të dhënave, është një aplikacion vizuel që përdorë vegla të ndryshme GUI(graphic user interface) për të transportuar të dhënat nga burimi në destinacion brenda warehouse. Shumica e aplikacioneve ETL duhet të qëndrojnë në sisteme operative GUI siç janë Windows apo Linux, përderisa disa mund edhe të punojnë në sistemin operativ UNIX. Kur kemi bazë të të dhënave që do të ruaj sasi të madhe të të dhënave, më së miri është që kjo bazë e të dhënave të vendoset në fushën gjenerale ku do të vendoset edhe data warehouse ashtu që të dhënat të mos kenë nevojë të lëvizin nëpër servera. Pasi që këto baza të të dhënave mund të jenë të mëdha, ato mund të kopjojnë shumë shpejtë shumicën e të dhënave nga bazat e të dhënave transaksionale, kjo duhet të mirret në konsideratë së bashku me hapësirën e nevojshme për data warehouse kryesor.

Duke filluar nga memorja, procesori dhe veçanërisht hapësira duhet të jenë në krye të listës së prioriteteve gjatë blerjes së harduerit për zhvillim. Kompanitë nuk duhet të marrin parasysh vetëm nevojat e sotme të bazës së të dhënave për instalim të projektit, por edhe çfarë përmirësimesh në data warehouse do të jenë të nevojshme në të ardhmen.

Kimball thekson: “shitësit më të mëdhenj të bazave të të dhënave relacionale janë ende platforma të zgjedhjes, duke mbuluar rreth 80% të tregut të DW/BI” [1]. Edhe pse ndoshta është e qartë se cilit sistem të bazës së të dhënave i referohet ai, është e domosdoshme të sigurohemi se këto sisteme mund të merren me madhësinë potenciale dhe nevojat e hapësirës që do të jenë të nevojshme, ndoshta në rangun terrabyte të shoqëruar me ndarje të arsyeshme të diskut. Gjatë planifikimit të hapësirës kompanitë gjithashtu duhet të mendojnë për planin për backup dhe rikthim të warehouse, veçanërisht backup fajllat mund të jenë shumë të mëdhenj.

Shtresa e qasjes në të dhëna kërkon një mjedis krejtësisht të ndryshëm. Edhe pse janë të mundshme teknologjitë e njehta të mbivendosjes, nuk mund të strehohen të gjitha në një lokacion, pasi sistemet dhe shtresa e data warehouse duhet të vendosen ndaras nga njëra tjetra. Në rastin e shtresës për qasje në të dhëna është e nevojshme një server për aplikacion. Për shtresën e të dhënave, “ka serverë që trajtojnë qasjen në të dhëna nga uebi, menaxhimi i pyetësorëve, raportimi në ndërmarrje, autentifikimi, bazat e të dhënave metadata, etj.” [1].

Pavarësisht kësaj, nevoja më e madhe për këtë shtresë i referohet pjesëve të uebit dhe aplikacionit që i japin shfrytëzuesve qasje të lehtë dhe të pastër në të dhënat e tyre nga veglat mbështetëse.

Mund të bëhet zgjedhja e platformës si aplikacioni SAP Business Objects i cili mund të veproj lehtësisht në Windows apo UNIX. Kompania duhet të vlerësojë se cila i përshtatet më së shumti situatës së kompanisë.

3.3 Metadata

Metadata përmbajnë informacionet dhe definicionet për të dhënat e ruajtura.

Elementet bazë të data warehouse janë dhënë në figurë:

Basic elements of the data warehouse

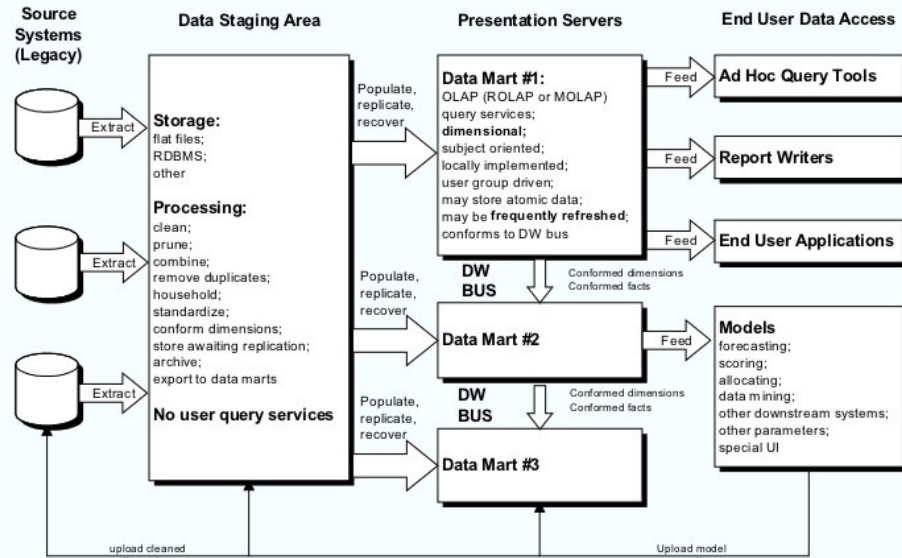


Figura 2. Elementet bazë të data warehouse

3.4 Tregu i të dhënave (Data marts)

Tregu i të dhënave është një nëngrup logjik i gjithë data warehouse dhe i përgatitur për një proces të vetëm biznesi në kompani. Kur këto bashkohen formohet një data warehouse i integruar e kompanisë. Tregu i të dhënave duhet të formohet nga faktet dhe dimensionet. Kështu ato mund të kombinohen dhe përdoren bashkërisht.

3.4.1 Çka është tregu i të dhënave

Tregu i të dhënave është model i cili paraqet strukturën e njëjtë të të dhënave me atë të data warehouse. Ato janë të përgatitura për kërkesa specifike të gjithë kompanisë apo një pjese të saj. Tregu i të dhënave i jep disa përparësi shfrytëzuesve pasi përmbanë më pak të dhëna. Së pari u mundëson të punojnë me pyetësorë më të shpejtë. Përparësi tjetër është lëvizshmëria pasi kërkon më pak hapësirë në disk që do i jep mundësi shfrytëzuesit të merr tregun e të dhënave në laptop. Gjatë procesit të dizajnit të tregut të të dhënave është e mundur të ndiqen dy metoda të ndryshme për të mbledhur të dhënat. Njëra zgjidhje është të mbledhen të dhënat nga data warehouse e kompanisë dhe pastaj t'i procesoj ato sipas nevojave mbi të cilat është përgatitur tregu i të dhënave,

zgjidhja tjetër është të mbledh të dhënat e formuara direkt në tregun e të dhënave. Të dhënat të cilat janë dizajnuar sipas kërkesave të tregut të të dhënave pastaj mbahen në një depo qendrore të të gjitha të dhënave të kompanisë. Këto zgjidhje mund të shihen në figurë.

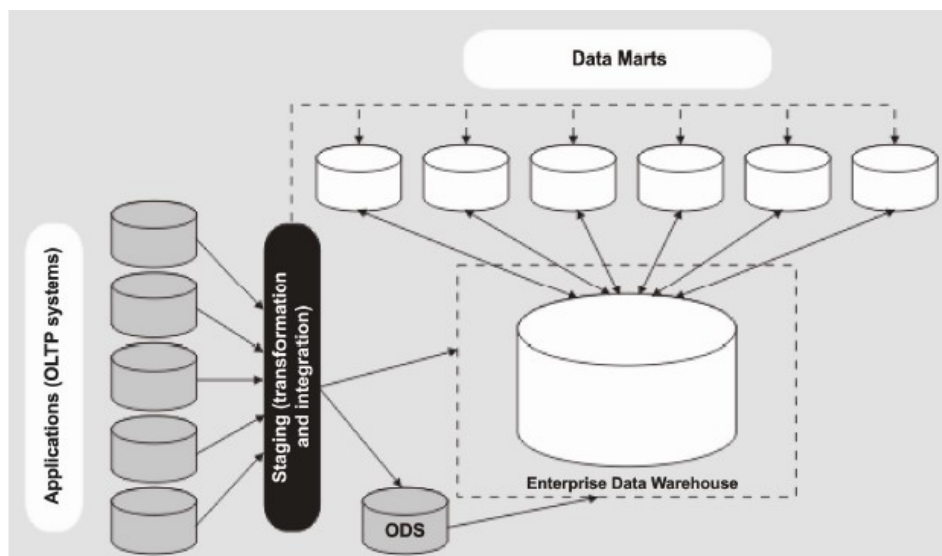


Figura 3. Arkitektura informative e ndërramrjes

Tregjet e të dhënave mund të kenë strukturë të varur apo të pavarur. Nëse karakteristikat e dimensioneve të tregut të të dhënave përcaktohen në fillim se si të përputhen njëra me tjetrën, atëherë ky treg i të dhënave do të ketë karakteristika të varura. Në disa raste është më mirë të kemi treg të të dhënave të pavarur. Karakteristikat e tregjeve tjera të të dhënave nuk do të mirren parasyshë gjatë përgatitjes së tregut të të dhënave. Megjithatë, kjo mund të pengoj integrimin në të ardhmen dhe do të shtoj koston e zhvillimit nëse do të ketë interesim në përhapjen e informatave ndërmjet departamenteve.

3.5 Shtigjet e zhvillimit të tregut të të dhënave

Për zhvillimin e tregut të të dhënave i kemi tri shtigje kryesore: shtegu lart-poshtë, shtegu poshtë-lart dhe shtegu i federuar.

3.5.1 Shtegu Lart-Poshtë

Sic shihet në figurën e mëposhtme të dhënat fillimisht vijnë në fushën data staging nga burimet operationale dhe në këtë fushë të dhënat kalojnë nëpër disa procese. Pas kësaj bëhet transferimi në data warehouse i cili pastaj i pason ato në tregun e varur të të dhënave.

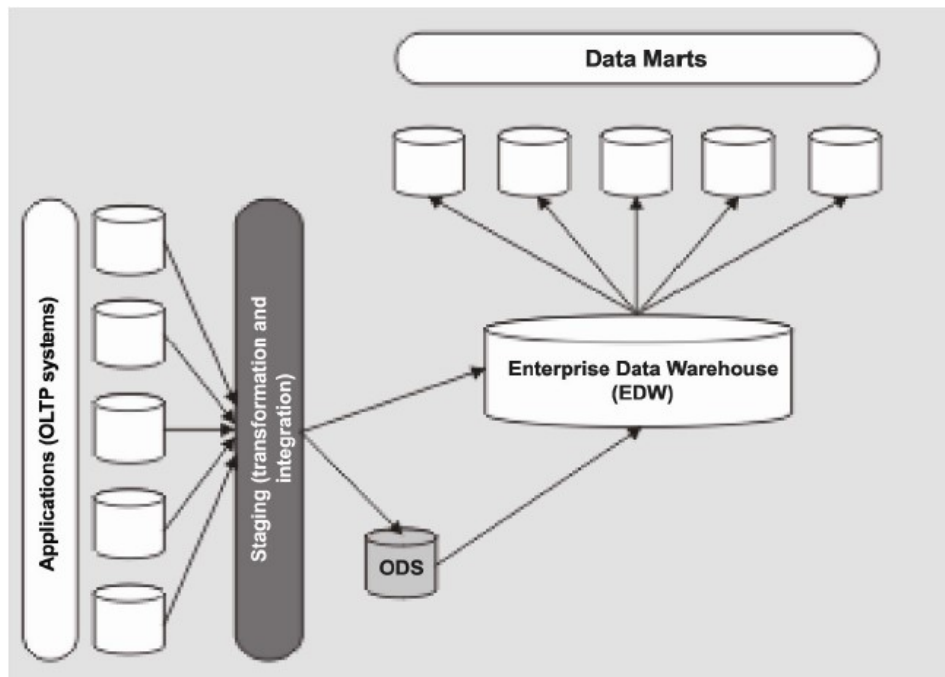


Figura 4. Shtegu lart - poshtë

3.5.2 Shtegu Poshtë-Lart

Në këtë shteg të dhënat të cilat vijnë nga sistemet trashëguese në fushën data staging, rrjedhin direkt në tregjet e të dhënave të pavarura dhe pastaj këto tregje të të dhënave mbushin data warehouse të kompanisë siç shihet edhe në figurë:

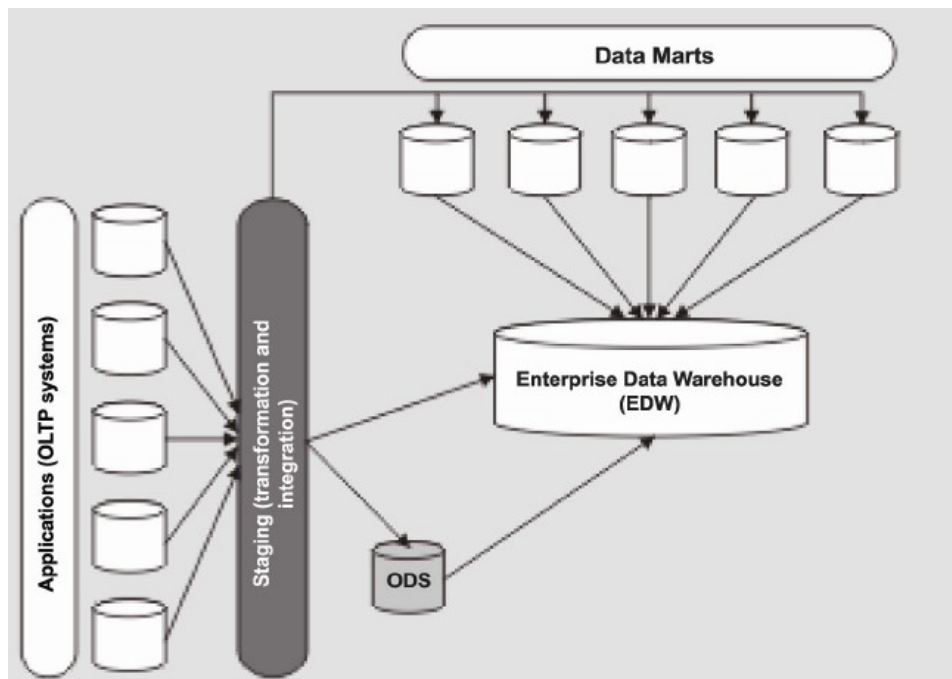


Figura 5. Shtegu poshtë – lart

3.5.3 Shtegu i Federuar

Me shtegun e federuar kemi të dy llojet e tregut të të dhënave edhe të varur edhe të pavarur, siç shihet në figurën 6. Tregu i pavarur i të dhënave i mbledh të dhënat direkt nga fusha data staging, kurse tregu i varur i të dhënave i merr të dhënat nga tregu i të dhënave.

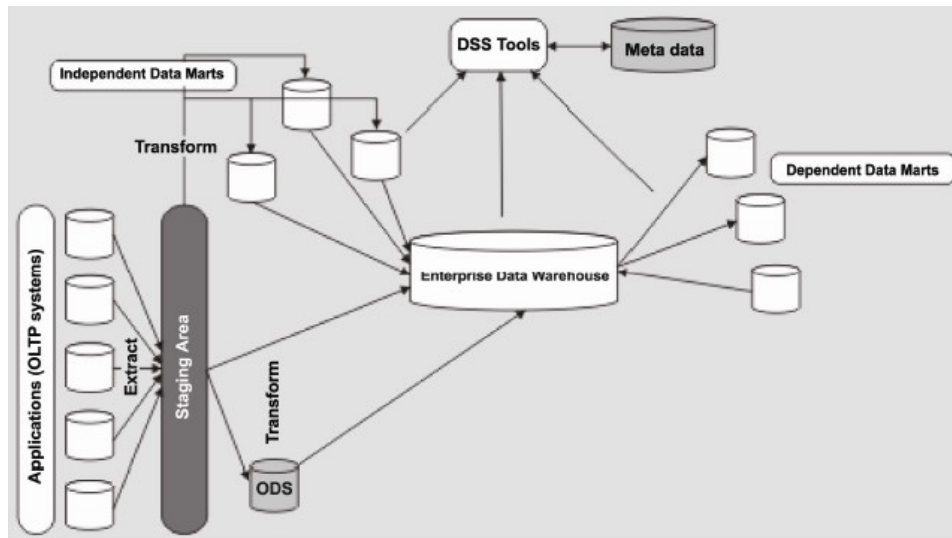


Figura 6. Shtegu i federuar

3.6 Dallimet ndërmjet data warehouse dhe tregut të të dhënave

Gjatë krahasimit të tregut të të dhënave me data warehouse mund të vërejmë lehtë dy dallime themelore. Njëra prej tyre është se tregu i të dhënave është një nëngrup i data warehouse dhe është i orientuar në kërkesa. Përkundër kësaj data warehouse përmbanë të dhëna të kompanisë pa u bazuar në kërkesa specifike. Por natyrisht, gjatë dizajnit të tregut të të dhënave duhet të mirret parasysh e gjithë struktura e data warehouse, nëse kjo nuk mirret parasysh atëherë do të jetë shumë i vështirë integrimi i mëvonshëm i tregut të të dhënave.

Implementimi i tregut të të dhënave është më i shpejtë dhe më i lirë, derisa ai përmbanë vetëm pjesë specifike të data warehouse, implementimi i së cilës merr më shumë kohë dhe ka kosto më të lartë.

Ka shumë zgjidhje të tregut të të dhënave që janë zhvilluar nga shitës të sistemeve mbështetëse për vendimmarrje. Por përdorimi i tyre për të dizajnuar një treg të të dhënave me kërkesa specifike kërkon përpjekje më të madhe për të rregulluar ato; për shkak se ato zgjidhje janë prodhuar për qëllime gjenerale.

Diferenca tjetër kryesore e tregut të të dhënave nga data warehouse është se të dhënat në tregun e të dhënave mund të jenë më të shpërndara se në data warehouse. Megjithatë kërkesat e tregut të të dhënave janë më të definuara se ato të data warehouse, grumbullimi i të dhënave mund të bëhet përgjatë kërkesave. Pra nxjerrja e të dhënave mund të bëhet në mënyrë më të shpejtë dhe më efikase.

3.7 Skema yll

Skema yll është pika qendrore e zhvillimit të data warehouse, pasi ajo siguron një strukturë të denormalizuar të bazës së të dhënave për raportim të të dhënave që janë të nevojshme për nevojat e kompanisë. Skema yll thjeshtë përkufizohet si “dizajn dimenzional për bazë të të dhënave relacionale“. Kjo skemë emrin e ka marrë nga pamja, që shfaqet si formë e yllit me tabelat dimenzion që rrethojnë pikën qendrore të skemës, që është tabela fakt. Çelësat zëvendësues janë çelësa primar të indeksuar në numra të plotë të tabelës dimenzion i cili bëhet çelës i huaj vetëm për lidhje me tabelën fakt, dhe kjo krijon bazën e skemës yll. Posa të vendosen lidhjet në tabelat dimenzion, ylli fillon të marrë formë, dhe kjo është shumë e zakonshme në data warehouse. Tabela fakt posedon shumë çelësa që lidhen me tabelat dimenzion dhe kështu mundësohet skema yll. Meqenëse tabela fakt përshkruan një proces të biznesit, “ajo duhet të përmbajë një grup gjithëpërfshirës të matjeve, edhe pse disa janë të tepërta“. Kjo siguron mundësinë për të bërë raportim më të lehtë duke i pasur të dhënat në dispozicion në atë tabelë.

Tabela më e ngjashme me atë të burimit të të dhënave dhe më e vlefshme në skemën yll është tabela dimenzion. Tabela dimenzion përmbanë pjesën më të madhe të të dhënave në data warehouse, dhe ka për detyrë të ruajë atributet e të dhënave për bartje të mëvonshme në tabelën fakt. Tabela dimenzion do të “përbëhet nga lidhja e fortë e një grumbulli të attributeve për të paraqitur objektet kyçe të biznesit, siç janë produktet, konsumatorët, punëtorët apo pajisjet“ [1]. Tabela dimenzion do të përmbajë psh. në një tabelë adresën e ndokujt dhe të gjithë emrat lidhur me të në një tabelë tjetër dimenzion. Tabelat dimenzion janë tabelat më të mëdha në data warehouse në krahasim me tabelat fakt dhe përmbajnë të dhëna raportuese nëse bëjmë pyetësorë në ta. Gjatë procesit ETL të dhënat së pari do të mbushen në tabelën dimenzion pasi ajo do të ketë çelësin primar që më pas do të referohet nga tabela fakt. Çelës primar i tabelës fakt do të jetë fusha e indeksuar nga tabela dimenzion e cila është me numra të plotë që rriten për një, e njohur si kolonë identitet në SQL Server. Kjo ID e tabelës dimenzion do të kopjohet në tabelën fakt për qëllime referuese gjatë procesit ETL.

3.8 Skema snowflake

Kemi edhe një skemë tjetër në data warehouse të njohur si skema snowflake, dhe kjo është një mënyrë tjetër për të lidhur të dhënat me skemat yll ekzistuese.

Kjo skemë do të shfaqet kur “lidhjet ndërmjet attributeve të dimensioneve bëhen të qarta në dizajnin dimensional“. Duke pasur të dhëna të ngjashme në disa baza të të dhënave apo në të njejtën bazë të të dhënave, është e mundur të ekzistojë një lidhje brenda skemave yll varësisht nga nevojat që do të lindin në kuadër të kërkesave të kompanisë.

Sikurse skema yll që e merr emrin nga forma që e merr gjatë dizajnit, edhe kjo skemë e merr emrin nga degëzimet që i bëhen skemës yll në tabelat dimensionale.

3.9 Mbushja dhe procesi ETL

Njëra nga pjesët më të rëndësishme të data warehouse është ekstraktimi, transformimi dhe mbushja e të dhënave nga bazat e të dhënave operative transaksionale në vetë data warehouse.

Edhe pse është e mundur të ndërtohet një ETL që do të kompletoj detyrën e saj ndërmjet bazës së të dhënave transaksionale dhe skemës yll të data warehouse, është më e logjikshme të ekstrahohen të dhënat dhe të ruhen përkohësisht para se të transformohen. Fusha e përqendrimit të të dhënave ekziston vetëm për të tërhequr të dhënat nga sistemi operativ, pastaj të riaranzhoj dhe pastroj të dhënat para se të ngarkohen në data warehouse.

Një data warehouse, e veçanërisht një e tillë për një ndërmarrje ka të dhëna që vijnë nga burime të ndryshme dhe kjo e bën ETL që të veproj më gjatë. Prandaj, tërheqja e shpejtë e të gjitha burimeve të të dhënave ka kuptim në këtë shkallë. Pasi të mirren të dhënat nga bazat e të dhënave transaksionale, duke i mbajtur të dhënat larg nga operacionet biznesore dhe në një vend të izoluar, ETL i ipet mundësia të transformoj të dhënat me lehtësi. Pasi të dhënat të kalojnë transformimin në mënyrë të arsyeshme, të dhënat pastaj kalojnë nëpër një hap të lehtë të ETL ku mbushen tabelat fakt dhe dimensionale të data warehouse.

Nuk ka analogji më të mirë të kësaj pjese të data warehouse se definicioni i Kimballit për fushën e përqendrimit të të dhënave të procesit ETL. Ai thotë se kjo krahasohet me kuzhinën e një restorani ku ushqimi i papërpunuar bëhet ushqim i mirë që konsumatorët do ta shijojnë. Kimballi e sheh krahasimin si: “në data warehouse, të dhënat operationale të papërpunuara transformohen në mënyrë të përshtatshme në warehouse për përodrim dhe pyetësorë nga shfrytëzuesit.“ [1] Duke marrë një hap më tej, hapi kur ushqimi bëhet i gatshëm, kuzhinieri i trajnuar do ta kompletoj detyrën e tij ashtu si një zhvillues i trajnuar i

BI do të krijoj një ETL për të vepruar njëjtë me të dhënat. Për më tepër, sikurse konsumatori i cili nuk ka gjasë të shohë kuzhinën se ku përpunohet ushqimi, ashtu edhe shfrytëzuesi nuk do të dijë për bazën e të dhënave që i përmbanë të dhënat.

Një nga arsyet kryesore që fusha e përqendrimit të të dhënave është e nevojshme të fillojë procesin e punës me sasi të madhe të të dhënave që krijojnë data warehouse final të organizuar. Në përkufizimin më të thjeshtë, arsyeja pse është e nevojshme kjo fushë është: për shkak të cikleve të ndryshme të kompanisë, kufizimet e harduerit dhe rrjetit dhe faktorët gjeografik, nuk është i mundur ekstraktimi i të gjitha të dhënave nga të gjitha bazat e të dhënave operacionale në kohë të njëjtë. Përderisa të dhënat vijnë nga burime të ndryshme dhe në forma të ndryshme, është e nevojshme që të ketë një vend për këto të dhëna në operacionet me shkallë të lartë të lëvizjes së të dhënave. Një shembull do të ishte që në sferën e data warehouse për kompaninë, të dhënat që vijnë nga sistemi i financave apo leximeve duhet të ekstrahohen shumë më shpesh se që do të duhej një bazë e vogël e të dhënave që ka të bëjë me sistemin e komunikimit.

Një tjetër përparësi e të pasurit të ruajtura të dhënat operacionale është mundësia për të ditur më shumë rreth të dhënave që do të ngarkohen në data warehouse për të përcaktuar cilësinë e të dhënave që shfrytëzuesit dëshirojnë të raportojnë në portalin e ri të tyre të BI. Aftësia për të marrë pjesë në profilizimin e të dhënave u jep teknikëve të data warehouse një grumbull standardesh që duhet ndjekur dhe të testojnë të dhënat që vijnë nga bazat e të dhënave transaksionale. “Profilizimi i të dhënave është analizimi i burimit të të dhënave për të kuptuar gjendjen e burimit të të dhënave“ [6], dhe i jep stafit të IT-së mundësinë për të analizuar të dhënat dhe të fillojnë rrugën për në data warehouse. Një përparësi e këtij hapi është se kur fillon faza e ekstraktimit të të dhënave, nëse numri i tyre është jonormalisht i madh apo i vogël, bëhen të mundura disa veprime nga ana e stafit për të përcaktuar nëse ka ndonjë gabim në ETL, apo duhet të ndryshojë ndonjë gjë në ETL ashtu që të veproj më me efikasitet. Mund të ketë edhe analiza statistikore që paraqesin përparësitë e të paturit një data warehouse në kompani, ashtu që menaxhmenti të shohë numrat përmbledhës që tregojnë investimet e tyre.

Me të dhënat e nxjerra në mënyrë të sigurtë nga bazat e të dhënave transaksionale, ndoshta pjesa më e favorshme e fushës së përqendrimit të të dhënave është mundësia e transformimit të të dhënave ashtu që raportimi në data

warehouse të jetë më i forcuar. Ndonëse të dhënat në bazat e të dhënave operacionale janë më efikase dhe padyshim më korrekte për aplikacionet që i mbështesin, mundësia e transformimit të të dhënave për data warehouse është një hap vendimtar dhe përparësia kryesore e operacioneve ditore të data warehouse. Përveç të dhënave të cilat janë tashmë nga bazat e të dhënave transaksionale, ekziston mundësia e riorganizimit të të dhënave në formë të tabelave të përkohshme, si dhe mënyra të tjera të transformimit të të dhënave të pa përpunuara para se të dërgohen në data warehouse. Një tjetër alternativë është trajtimi i temave tabu të standardizimit të të dhënave që nuk është i mundur në sistemet e mëparshme të aplikacioneve dhe të cilat kanë krijuar zakone të këqija të shfrytëzuesit. Kështu, “standardizimi dhe koordinimi i të dhënave është një grup i proceseve që kryesisht përbëhet nga dizajnimi dhe ekzekutimi, puna e standardizimit është të krijoj njëtrajshmëri ndërmjet elementeve specifike të të dhënave“ [6].

Ekzistojnë mundësi të mrekullueshme kur bëhet fjalë për transformimin e të dhënave, të cilat në të kaluarën nuk kanë mundur të ndryshohen. Kimball thekson se: “ka transformime të shumta të mundshme, të tilla si pastrimi i të dhënave (korigjimi i gabimeve drejtshkrimore, zgjidhja e konflikteve të domenit, apo parsimi në format standarde), kombinimi i të dhënave nga burime të ndryshme, dhe caktimi i çelësave në warehouse“ [1].

Duke pasur të dhëna nga burime të ndryshme por për një destinacion të vetëm kërkon ndërhyrjen e ETL për të kthyer të dhënat konform standardeve. Me transformime tjera në dispozicion sikurse vegla SQL Server Integration Service ETL, ekziston mundësia e shtimit të kolonave bazuar në derivimin e të dhënave.

Fusha e përqendrimit të të dhënave është njëra nga pjesët më të rëndësishme të data warehouse pasi ofron mundësinë e organizimit të të dhënave nga të gjitha burimet në një vend para se të dërgoj të dhënat në destinacionin final ku shfrytëzuesit kanë mundësinë të bëjnë pyetësorët. Sikurse analogjia e Kimballit me kuzhinën e restorantit, kjo fushë nuk do të jetë transparente për shfrytëzuesit përveç zhvilluesve përgjegjës për ETL. Edhe pse nuk është e nevojshme, duket të jetë vendim i drejtë të vendosen të dhënat para se të mbushet data warehouse, veçanërisht me një sistem më të madh apo fushëveprim të data warehouse. Përdërisa kemi të dhënat e vendosura në një vend, i jep mundësinë ETL të kompletoj hapin e mesëm të transformimit të të dhënave dhe të mënjanoj ndonjë

anomali gjatë rrugës. Pas kompletimit, hapi tjetër është mbushja me të dhëna e tabelave fakt dhe dimenzion.

3.10 Alternativa ETL

Ka shumë mënyra për të ndërtuar një skemë yll, por tradicionalisht rruga më e mirë është të dizajnojmë skemën bazuar në nevojat e kompanisë. Megjithatë, ka disa raste kur skema yll mund të gjenerohet bazuar në të dhënat hyrëse. Tre akademikët kanë shpikur një mënyrë gjithëpërfshirëse për të gjeneruar skemën e data warehouse duke u bazuar në bazën e të dhënave transaksionale dhe programin e zhvilluar për vlerësim të të dhënave. Në këtë rast, ata dallojnë propozimin e tyre duke pohuar se “puna nuk kufizohet në dizajnimin e skemës por ajo gjithashtu popullon të dhënat nga burimet e grumbulluara të të dhënave në tabela të skemave të gjeneruara automatikisht” [7]. Ata përdorin grumbullimin hierarkik për të ndihmuar gjenerimin jo vetëm të skemës yll por edhe të skemave tjera. Programi i tyre gjeneron skemën dhe pastaj vendos produktin final në serverët e bazave të të dhënave. Një program i tillë është një përmirësim i madh mbi modelimin e zakonshëm dhe zhvillimin manual të skemës yll e pastaj dizajnimi dhe programimi i ETL për të futur të dhënat në data warehouse. Modeli i propozuar i tyre për skema automatike është i përbërë nga disa pjesë për të përfunduar procesin e marrjes së të dhënave të pa përpunuara në skemat eventuale në data warehouse. “Komponenti gjithashtu trajton parametrat e grumbulluar të cilët përfshijnë metodën e lidhjes (i vetëm, komplet apo mesatar), drejtimet e grumbulluara (rreshtat, kolonat, të dyjat) dhe matjen distancë/ngjashmëri (përputhje Euklidiane, Menhatane, Të thjeshtë për koeficientin korelacional të Pearson-it)” [7]. Kjo është e rëndësishme pasi grumbullimi hierarkik është përbërës në hapin e ardhshëm që përdoret për të degjeneruar skemën. Në shtresën tjetër, të dhënat e grumbulluara bëhen pjesë e gjenerimit të skemës duke lexuar në të dhënat e grumbulluara në hapin e mëparshëm. Pastaj, përcaktohet lloji i tabelës i cili është fakt apo dimenzion dhe gjithashtu përcaktohen lidhjet ndërmjet tabelave. Dhe në fund, të dhënat kalojnë në tabelat finale që përbëjnë skemën e dëshiruar, qoftë ajo yll, snowflake apo galaktikë.

Autorët pastaj demonstrojnë përdorimin e mjetit të tyre duke përdorur të dhëna nga një tekst fajll që përfundojnë në skemën yll. Ata tregojnë se për eksperiment kanë përdorur algoritmin Euklidian, dhe “arsyeja për këtë zgjedhje është se distanca Euklidiane është se kjo distancë përdoret më së shumti në analiza” [7]. Programi të cilin e përdorin për të ekzekutuar procesin është një

program bazik C#, i cili i ka ndjekur hapat e nevojshëm për të krijuar një skemë yll për këtë eksperiment. Është interesante të mendohet se kjo metodologji që përdorë një algoritëm do të jetë më efektive se dizajnimi, ndërtimi dhe ngarkimi manual i skemës nëpërmjet ETL. Megjithatë, është diskutabile nëse ky lloj programi është i përshtatshëm për të gjitha situatat apo do të jetë unik sikurse ETL, mirëpo padyshim se nëse ky program do të përdoret në mënyrë univerzale do të kursente shumë kohë. Në data warehouse të ndërmarrjeve ku të dhënat hyrëse vijnë nga të gjithë sistemet e ndërmarrjes, ndoshta një aplikim individual për secilën hyrje do të ishte e arsyeshme.

3.11 Biznis Inteligjenca

Edhe pse shumica e bazave të të dhënave që përdoren në kompani kanë nga një aplikacion për përdorim më të lehtë të tyre, data warehouse është vetëm një bazë e madhe e të dhënave që nuk ka një shtresë të aplikacionit. Megjithatë ka aplikacione që janë pjesë e data warehouse, por ato janë aty për të plotësuar depon si komponentë që ofron qëllime të ndryshme. Së pari, aty kemi një aplikacion ETL(Extract, Transform, Load) që është pjesë e zgjidhjes që bartë të dhënat nga baza e të dhënave burimore e transaksioneve. Aplikacionet më të njohura ETL janë: SQL Server Integration Services, OBIEE nga 'Oracle and Informatica'. Pjesa më e njohur e data warehouse janë veglat për raportime që përodren si pjesë e shtresës së prezentimit. Aplikacioni i raporteve nuk është pjesë e data warehouse, por është pjesa që paraqet data warehouse dhe me të cilën shfrytëzuesit janë të familjarizuar. Ka një mori veglash për të nxjerrë të dhënat nga data warehouse, të cilat janë zakonisht homogjene me data warehouse ashtu që mund të përodren me baza të tjera të të dhënave për raportime apo prezentim të të dhënave. Megjithatë, pjesa më e rëndësishme është aplikacioni i cili përmbanë pjesën e prezentimit, e cila është komponentë e përhapjes së informacionit, e veçantë e raportimit.

3.12 Raportimi

Tani më shumë se kurrë më parë, kompanitë mbështeten në informata të sakta, të sigurta dhe më të freskëta në raportet e tyre për të ndihmuar operacionet e tyre të biznesit dhe në përputhje me rregulloret gjithnjë e më të rrepta.

Çështja është se mund të ketë të dhëna në lokacione të ndryshme, dhe ato jashtë IT-së mund të nxjerrin raporte pa marrë parasyshë saktësinë e të dhënave. Rastet e tilla mund të parashtrojnë pyetje në lidhje me besueshmërinë e të dhënave në raport dhe mund të dërgojnë në vendime jo të mira dhe çështje

ndoshta edhe më të mëdha në vazhdim. Në kontekstin e të pasurit një data warehouse në kompani, veglat BI sigurojnë shpërndarjen në kohë të të dhënave për raportim dhe analizim nëpër ndërmarrje. Përveç kësaj, mjetet të cilat menaxhojnë, konsolidojnë dhe shpërndajnë të dhënat gjithashtu përmbajnë një gjurmë të auditimit të të dhënave, prej nga vijnë, çka është bërë me to, dhe kush i pranon ato. Një zgjidhje efektive e BI duhet të siguroj shpërndarjen në kohë të të dhënave, të lejoj raportimin dhe analizimin nëpër ndërmarrje, të ofroj vegla lehtë të përdorshme për menaxhim, konsolidim dhe shpërndarje të të dhënave dhe më e rëndësishmja është të përmbajë gjurmë të auditimit të të dhënave. Pika kulmore e data warehouse është se me çfarë efektiviteti zgjidhja e BI ndërvepron me të dhënat. Veglat më të njohura për raportim janë SAP Business Objects, IBM Cognos, Jaspersoft dhe Pentaho.

Arsyet për të marrë parasysh se kur duhet blerë e kur duhet zhvilluar raportet përfshijnë: numrin total të raporteve, shpërndarjen e dëshiruar dhe aftësitë ad-hoc. Me veglat superiore në dispozicion, pyetja me të cilën ballafaqohen bizneset gjatë implementimit ende është: a duhet blerë apo zhvilluar vegla për raportim për nevojat tuaja të BI.

Kur vendoset se cilën vegël BI duhet përdorur, për shkak të investimeve të konsiderueshme dhe shqetësimit në lidhje me përdorshmërinë, menaxhmenti duhet të përcaktoj llojin më të mirë të produktit për ta zbatuar. Nivelet e modelit rriten bazuar në produktet tradicionale dhe standarde të BI që janë implementuar.

Në “Magic Quadrant for Business Intelligence Platforms“ nga Gartner Research, kuadranti paraqet ku qëndrojnë veglat BI në krahasim me të tjerat kur analizohet ndikimi i veglës BI.

Kuadranti përfshinë lojtarët e ngrohtë, sfiduesit, udhëheqësit dhe vizionarët. Boshti vertikal paraqet aftësinë për të ekzekutuar, ndërsa boshti horizontal paraqet plotësinë e vizionit. Liderët dhe lojtarët e ngrohtë kanë më së shumti grupe, pasi liderët janë më të njohurit, ndërsa lojtarët e ngrohtë janë pjesë e një nevoje të caktuar të biznesit. Artikulli harxhon kohën duke kërkuar në trendet e tregut se çfarë kërkojnë shfrytëzuesit nga veglat BI dhe ku janë preferencat e tyre këto ditë. Ata ecin nëpër çdo produkt, duke përfshirë veglat më të njohura të kompanive të mëdha, si SAP, Microsoft dhe Oracle, dhe diskutojnë për pikat e forta dhe dobësitë që lidhen me çdo vegël.

Pavarësisht aplikacioneve që janë sinonim me BI që kanë të bëjnë me data warehouse, vegla e parë e cila përdoret në shtresën e prezentimit është vegla bazë për pyetësorë. Vegla për pyetësorë “i lejon shfrytëzuesit të krijoj pyetësorë direkt në modelin dimensional dhe të përcaktoj një grup rezultatesh“ [1]. Këtë vegël zhvilluesit e BI do ta përdorin edhe për të zhvilluar pyetësorë për të zhvilluar raporte të reja apo zgjidhje të data mining.

“Qëllimi i gjenerimit të raporteve është të transformojnë të dhënat në informata për shfrytëzuesit e biznesit“ [8]. Vegla më e popullarizuar do të jetë ende raportimi ad-hoc me të cilën shumica e shfrytëzuesve janë të familjarizuar. Veglat më të popullarizuara të raportimit mbeten SAP Crystal Reporting Tool dhe Report Builder nga Microsoft. Vegla e Microsoft-it Report Builder është shtuar në SQL Server 2008. Ashtu si drejtohen raportet Crystal në Business Object Enterprise, raportet e Microsoft-it janë zhvilluar në SQL Server Reporting Service, të cilat veprojnë me të njëjtin kapacitet si vegla ueb për shpërndarjen e raporteve të shfrytëzuesit.

Këto magazina të raporteve i japin departamentit të IT-së mundësinë e qeverisjes me qasje në raportet përfshirëse në data warehouse, dhe ata i japin gjithashtu shfrytëzuesve mundësinë të tërheqin të dhënat në vend se të gjenerojnë raporte për ta si në të kaluarën. Një nga tendencat më të popullarizuara ditëve të sodit që merr avantazh në implementimin e data warehouse është përdorimi i ‘dashboards’, i cili i ofron mundësinë shfrytëzuesit për ndërveprim gjatë marrjes së informatës. “Dashboard-i siguron mjetet e ndërfaqeve që mbështesin raportimin e statuseve dhe alarmeve nëpër burime të ndryshme të të dhënave në nivele të larta, dhe gjithashtu lejon kërkimin më të detajuar të të dhënave“ [1]. Një dashboard tipik zakonisht përfshinë datasetet, grafet, treguesit kyç të informatës dhe të dhënat përfshirëse. Një nga përparësitë e përdorimit të dashboard është se në shumë raste do të jetë interaktive me shfrytëzuesit dhe është në gjendje të paraqes projeksione për t’iu mundësuar shfrytëzuesve të marrin vendime më të mira. Dashboard-i ka synim grupet e vogla apo audiencën e targetuar, pasi janë të orientuara në detaje të hapësirave specifike. Universiteti i Floridës shpjegon kalimin e tyre nga sistemi i vjetër i raportimit në dashboard në ndërthurje me data warehouse të tyre: “Ky ishte një ndryshim thelbësor. Para kësaj, ne kemi ofruar tipin e planifikimeve të të ardhurave tona në raporte me format spreadsheet. Dhe kjo tërësisht ka kaluar në sistem online ku njerëzit mund t’iu qasen të dhënave që u nevojiten dhe mund të krijojnë pyetësorë për të marrë të dhënat në formatin e dëshiruar.“ [9]

Tjetër përdorim i popullarizuar i BI në mjedisin e data warehouse është mundësia për të përdorur OLAP (Online Analytical Processing) të avansuar dhe data mining. OLAP mund të organizoj të dhënat në struktura shumë dimensionale për t'iu mundësuar shfrytëzuesve të zgjedhin dimensionin që dëshirojnë ta shohin.

Data mining do të “eksploroj të dhënat në data warehouse për modele dhe marrëdhënie të dobishme“ [1]. Vegla e popullarizuar për këtë është SQL Server Analysis Services, e cila është vegël e fuqishme analitike e përfshirë në kompletin e produkteve të BI. Vegla i jep mundësi shfrytëzuesit të dizajnoj kubet bazuar në skemën yll të data warehouse, pastaj shton matjet dhe indikatorët kyç të performancës të njohura si KPI(Key Performance Indicators).

Të gjithë shembujt e raportimeve i japin shfrytëzuesit mundësinë të konsolidoj të gjitha të dhënat në data warehouse dhe të krijoj një version të vetëm, i cili është thelbësor në saktësinë e raporteve. Kemi shumë vegla në dispozicion për përdorim në sistemin BI dhe ka burime për të gjetur se cili punon më mirë për kompaninë. Pasi shumë nga veglat DW/BI janë të ngjashme për nga natyra dhe fushëveprimi, më e rëndësishme për shfrytëzuesit është se i përshtaten mirë biznesit. Në përgjithësi, Kimball konkludon se veglat e BI duhet të jenë praktike, me përmbajtje të pasur, të pastra, aktuale, dhe interaktive dhe të orientuara në vlera.

Duke pasur një data warehouse, kompania mund të përfitojë nga teknologjitë më të reja që janë në dispozicion në këtë fushë dhe të fuqizojë shfrytëzuesit për të marrë vendime më të mira me informacionet që i kanë në dispozicion.

3.13 Implementimi

Ky kapitull përfshinë një shembull të implementimit të data warehouse në kompani, në një shkallë të vogël që përmbanë vetëm një skemë ‘snowflake’. Zgjidhja e paraqitur këtu është një shembull që përdorë veglat e biznes inteligjencës me SQL Server dhe bazë të të dhënave në SQL Server, i njëjti softuer që është i përshtatshëm edhe për implementim të tërësishëm.

3.14 Prototipi i data warehouse së ndërmarrjes

Kulminacioni i tezës vjen së bashku me prototipin e data warehouse në kompaninë KEDS. Qëllimi i këtij dizajni fillestar të data warehouse është të paraqes jo vetëm se si një seri e tabelave mund të bëhen një skemë ‘snowflake’ e

denormalizuar, por gjithashtu të paraqes se si data warehouse në kompani mund të integrojë bashkë të dhënat ndërmjet sistemeve të ndryshme.

Hapi i ardhshëm është marrja e të dhënave nga baza e të dhënave duke i ekstraktuar ato, pastaj të transformohen për t'i vendosur në grupe logjike, dhe në fund të ngarkohen në skemë 'snowflake' të organizuar. Për t'iu paraqitur shfrytëzuesit se çka është e mundur, kemi disa shembuj të raporteve që i paraqesin të dhënat. Raporti paraqet mundësitë që mund t'i ketë shfrytëzuesi në një raport. Të dhënat që fillimisht ishin në shtatë baza të të dhënave të ndryshme, kanë mundësi të lidhen në data warehouse. Ashtu siç e paraqet prototipi, mundësia për t'i lidhur të dhënat mund të ndihmojë në situatat kur nevojitet ndonjë raport krahasues ndërmjet departamenteve të ndryshme. Ky është një shembull i thjeshtë i krijimit të data warehouse, por shfaq konceptin e një pasqyre të madhe të implementimit të data warehouse. Data warehouse finale është baza e të dhënave në SQL Server 2012, që përdorë Server Integration Services si vegël për ETL, Server Analysis Services si vegël për krijimin e kubeve dhe për përgatitje të të dhënave për raportim dhe SQL Server Reporting Services si vegël e raportimit.

Tabelat dimenzion do të jenë direkt nga tabelat nga baza e të dhënave, duke i shtuar një dimension ID që është fushë identifikuese dhe rritet për një. Disa nga tabelat dimenzion do të kenë të dhëna edhe nga tabela të tjera, siç janë tabelat lookup. Tabela fakt do të ketë një fakt ID unike që lidhet me tabelat dimenzion. Në këtë shembull të gjitha tabelat fillimisht do të jenë të zbrazëta. Megjithatë, për shkak të qëllimit dhe madhësisë, ky data warehouse nuk do të përfshijë të gjitha detajet e data warehouse, mirëpo ende përcjell dizajnin bazë dhe rezultatet.

3.15 Konstruktimi i ETL

Në figurën e mëposhtme është paraqitur diagrami i bazës së të dhënave të data warehouse.

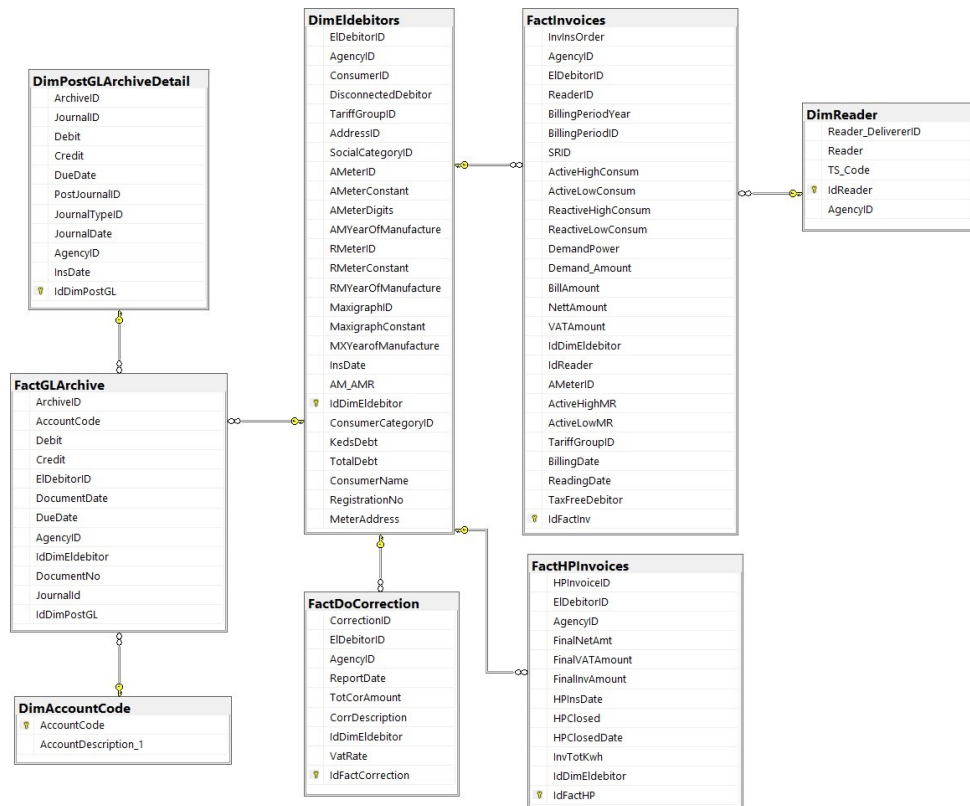


Figura 7. Diagrami i bazës së të dhënave të data warehouse

3.16 Mbushja e tabelave

Figurat në vazhdim tregojnë mbushjen e tabelave dimension dhe fakt. Së pari do të mbushen tabelat dimension dhe pastaj tabelat fakt. Tabelat fakt mbushen pas tabelave dimension për shkak të lidhjes me çelësat që gjenden në tabelat dimension.

Mbrenda Data Flow do të ndodhin shumë ekstrakte, transformime dhe ngarkime në të njëjtën kohë. Derisa të dhënat ekstrahohen nga shumë tabela, e gjithë puna e ETL mund të bëhet njëkohësisht. Aktivitetet mbrenda një tasku fillojnë me ekstrakt dhe do të zgjedhin tabelën e caktuar në SQL. Nëse është e nevojshme të lidhet me ndonjë tabelë tjetër, mund të përdoret tasku 'lookup' për të marrë të dhëna nga një tabelë tjetër para se të bëhet ngarkimi në tabelën destinuese.

3.16.1 Mbushja e tabelave dimension

Tabela dimension 'AccountCode'

Pyetësi i përdorur për të mbushur këtë tabelë është:

```
SELECT tcc.AccountCode,tcc.AccountDescription_1  
FROM tblCCGLAccountCode tcc
```

Paketa në SSIS për tabelën dimension 'AccountCode'

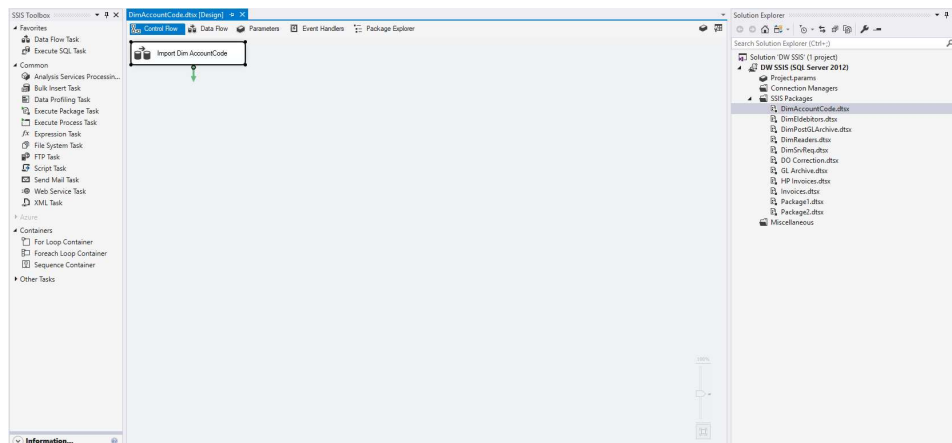


Figura 8. Control Flow i tabelës dimension 'AccountCode'

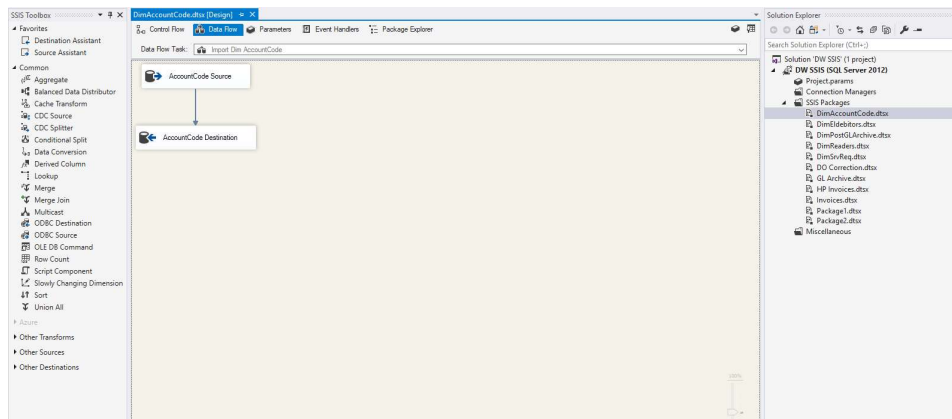


Figura 9. Data Flow i tabelës dimension 'AccountCode'

Tabela dimension 'ElDebitors'

Pyetësi i përdorur për të mbushur këtë tabelë është:

```
SELECT de.ElDebtorID, de.AgencyID, de.ConsumerID, de.DisconnectedDebtor,  
de.TariffGroupID, de.AddressID, de.SocialCategoryID, de.AMeterID, CAST  
(de.AMeterConstant AS INT) AS AMeterConstant, de.AMeterDigits,
```

```

de.AMYearOfManufacture, de.RMeterID, CAST(de.RMeterConstant AS INT) as
RMeterConstant, de.RMYearOfManufacture, de.MaxigraphID,
CAST(de.MaxigraphConstant AS INT) AS MaxigraphConstant,
de.MXYearofManufacture, de.InsDate, de.AM_AMR, de.ConsumerCategoryID,
(ISNULL(c.Company + ' ', '') + ISNULL(c.ConsumerFirstName + ' ', '') +
ISNULL(c.ConsumerMidleName + ' ', '') + ISNULL(c.ConsumerSecondName,
'')) AS ConsumerName, c.RegistrationNo,
(ISNULL(LTRIM(RTRIM(s.StreetName)) + ' ', '') +
ISNULL(LTRIM(RTRIM(de.AddressNo)), '')) AS MeterAddress

FROM tblCCEldebtors de left outer join tblCCConsumer c
ON de.ConsumerId=c.ConsumerId

LEFT OUTER JOIN tblCCStreets s ON de.StreetID = s.StreetID

```

Paketa në SSIS për tabelën dimension 'ElDebtors'

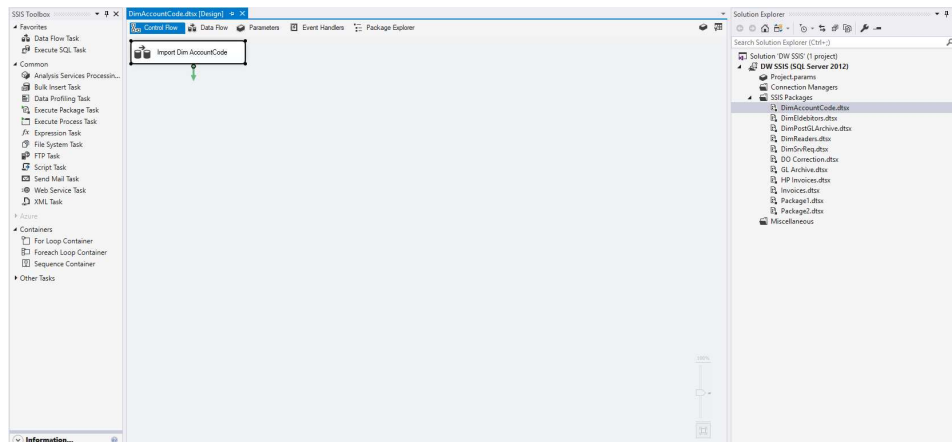


Figura 10. Control Flow i tabelës dimension 'ElDebtors'

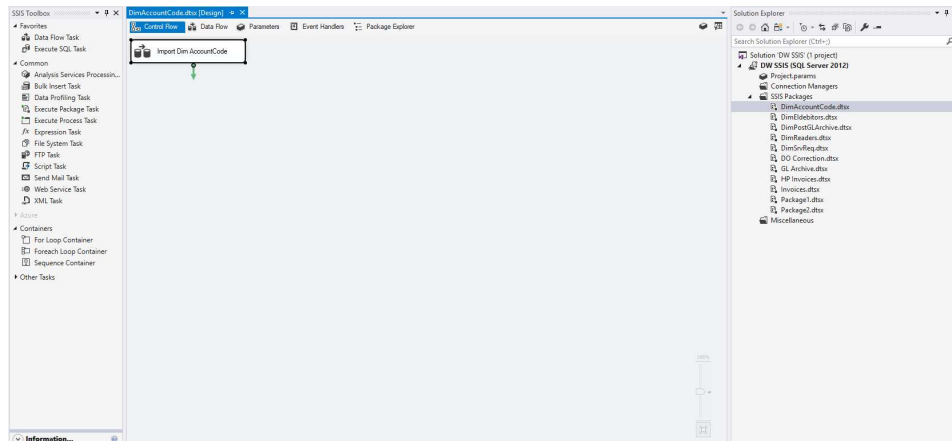


Figura 11. Data Flow i tabelës dimension 'ElDebitors'

Tabela dimension 'PostGLArchive'

Pyetësi i përdorur për të mbushur këtë tabelë është:

```
SELECT p.ArchiveId, p.JournalId, p.Debit, p.Credit, p.DueDate, p.PostJournalId,
j.JournalTypeId, j.JournalDate, LEFT(p.JournalId, 3) AS AgencyID, j.InsDate
FROM tblPostGLArchiveDetail p INNER JOIN tblPostGL_Journal j
ON p.PostJournalID = j.JournalID
WHERE YEAR(j.JournalDate) > 2016 AND j.JournalTypeId IN
('MC','MB','SR')
```

Paketa në SSIS për tabelën dimension 'PostGLArchive'

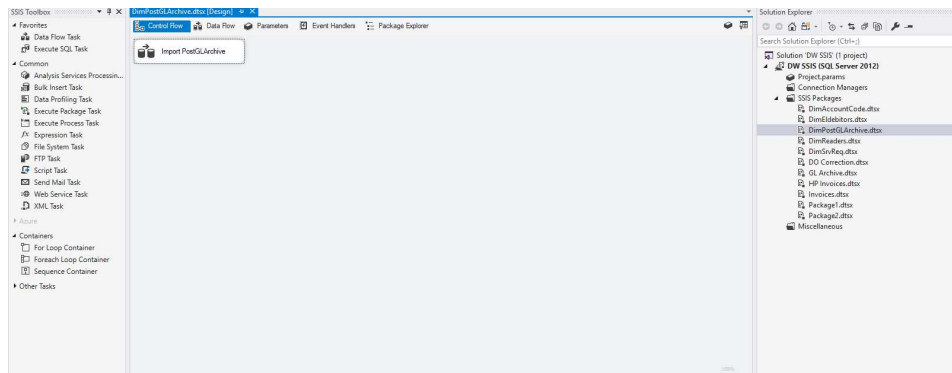


Figura 12. Control Flow i tabelës dimension 'PostGLArchive'

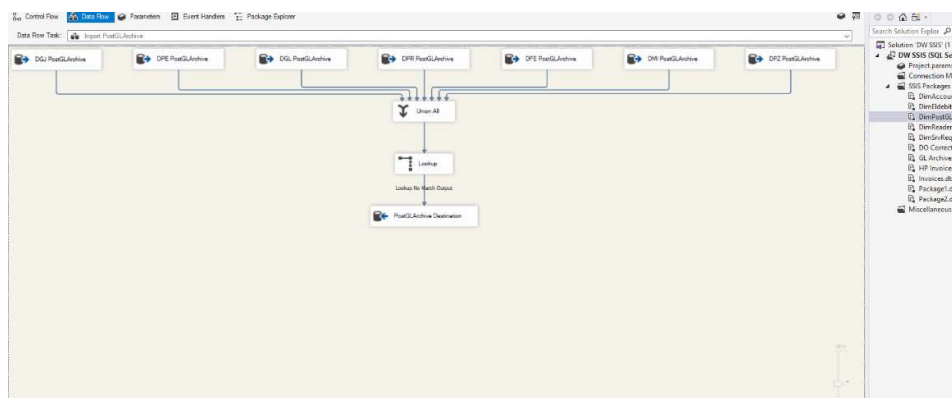


Figura 13. Data Flow i tabelës dimension 'PostGLArchive'

Tabela dimension 'Readers'

Pyetësi i përdorur për të mbushur këtë tabelë është:

```
SELECT AgencyId, tc.Reader_DelivererID, CAST (tc.Reader as nvarchar(50))
AS Reader,tc.TS_Code
```

```
FROM tblCCReader tc
```

Paketa në SSIS për tabelën dimension 'Readers'

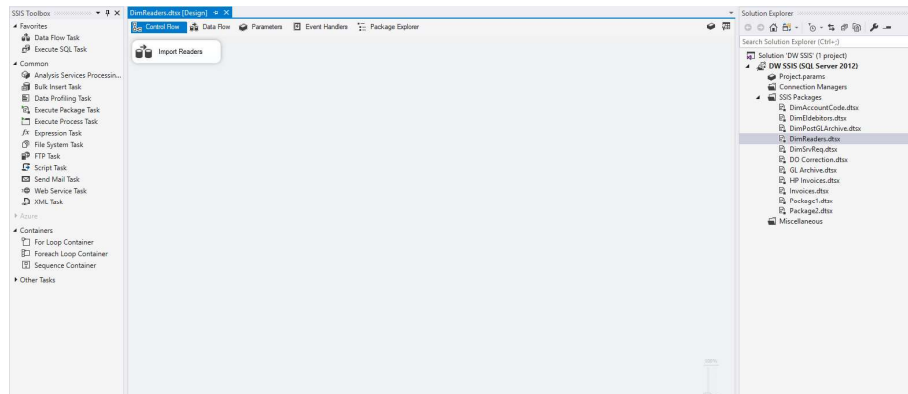


Figura 14. Control Flow i tabelës dimension 'Readers'

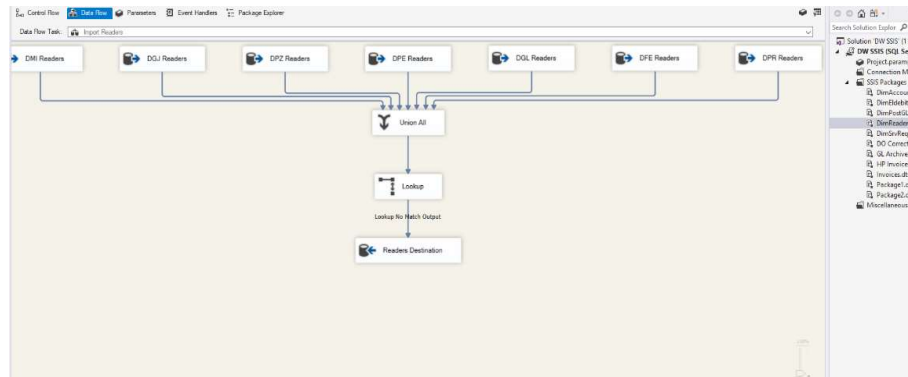


Figura 15. Data Flow i tabelës dimension 'Readers'

3.16.2 Mbushja e tabelave fakt

Tabela fakt 'Invoices'

Pyetësi i përdorur për të mbushur këtë tabelë është:

```
SELECT  tcb.InvInsOrder, tcb.AgencyID, tcb.EIDDebitorID, tcb.ReaderID,
        tcb.BillingPeriodYear, tcb.BillingPeriodID, tcb.SRID, tcb.ActiveHighConsum,
        tcb.ActiveLowConsum, tcb.ReactiveHighConsum, tcb.ReactiveLowConsum,
        tcb.DemandPower, tcb.Demand_Amount, tcb.BillAmount, tcb.NettAmount,
        tcb.VATAmount, tcb.AMeterID, tcb.ActiveHighMR, tcb.ActiveLowMR,
        tcb.TariffGroupID, tcb.BillingDate, tcb.ReadingDate, tcb.TaxFreeDebitor
```

```
FROM tblCCIInvoicesH tcb
```

```
WHERE BillingPeriodYear > 2016
```


Pas mbushjes së tabelës me të dhëna nga baza e të dhënave transaksionale, do të bëhet plotësimi i shënimeve nga tabelat dimension përmes stored procedurave:

exec updateFactInvoicesDebitor – për plotësim nga tabela dimension ‘ElDebitors’

exec updateFactInvoiceReader – për plotësim nga tabela dimension ‘Readers’

Paketa në SSIS për tabelën fakt ‘Invoices’

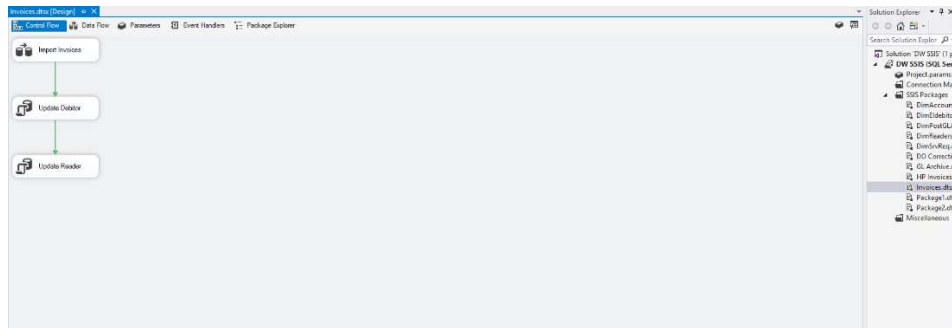


Figura 16. Control Flow i tabelës fakt 'Invoices'

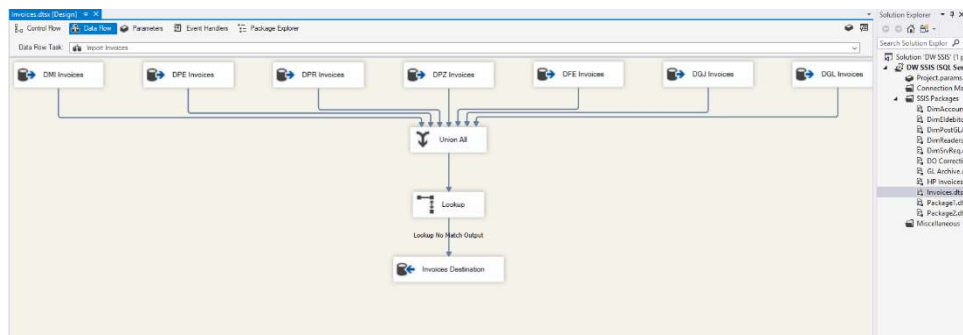


Figura 17. Data Flow i tabelës fakt 'Invoices'

Tabela fakt ‘HPInvoices’

Pyetëtori i përdorur për të mbushur këtë tabelë është:

```
SELECT th.HPInvoiceID, th.ElDebitorID, th.AgencyID, th.FinalNetAmt,
th.FinalVATAmount, th.FinalInvAmount, th.HPInsDate, th.HPClosed,
th.HPClosedDate, th.InvTotKwh
```

FROM tblHPInvoices tb

WHERE YEAR(tb.HPClosedDate) > 2015

Tabela fakt pastaj do të plotësohet me shënimet nga tabela dimension përmes stored procedurës:

exec updateFactHPInvoicesDebitor – për plotësim nga tabela dimension ‘ElDebitors’

Paketa në SSIS për tabelën fakt ‘HPInvoices’

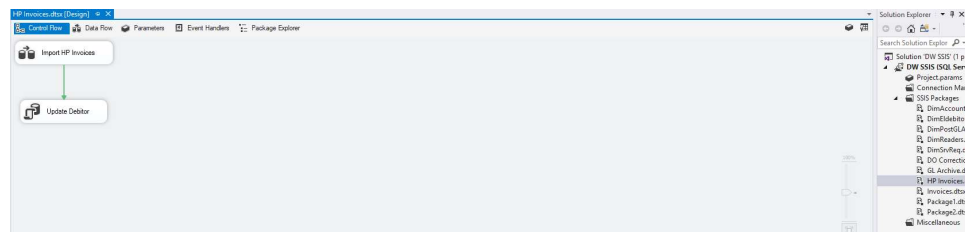


Figura 18. Control Flow i tabelës fakt 'HPInvoices'

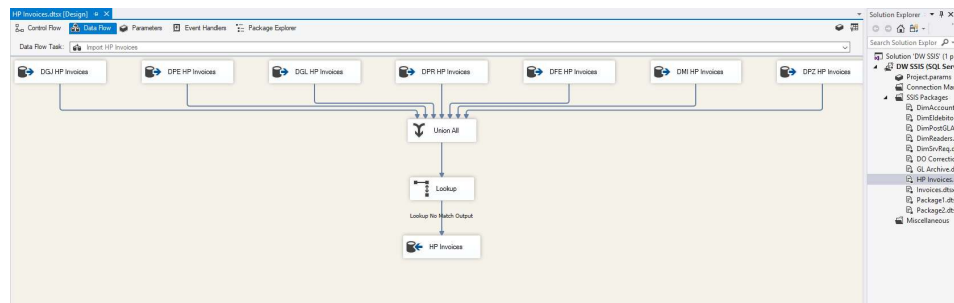


Figura 19. Data Flow i tabelës fakt 'HPInvoices'

Tabela fakt ‘GLArchive’

Pyetësi i përdorur për të mbushur këtë tabelë është:

SELECT 'DPE' AS AgencyID, tc.ArchiveID, tc.AccountCode, tc.Debit,
tc.Credit, tc.ElDebitorID, tc.DocumentDate, tc.DueDate, tc.DocumentNo, tc.JournalID

FROM tblCCGLArchive tc

WHERE YEAR(tc.Documentdate) > 2016

Pas mbushjes së tabelës fakt me shënime, do të plotësohen shënimet nga tabela dimension përmes stored procedurave:

exec updateFactGLArchiveDebitor – për plotësimin e shënimeve nga tabela dimension ‘ElDebitors’

exec updateFactGLArchivePostGL – për plotësimin e shënimeve nga tabela dimension ‘PostGLArchiveDetail’

Paketa në SSIS për tabelën fakt ‘GLArchive’

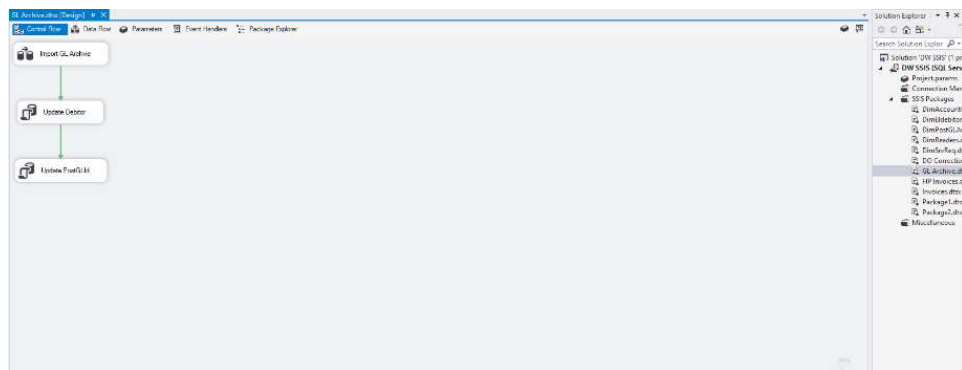


Figura 20. Control Flow i tabelës fakt 'GLArchive'

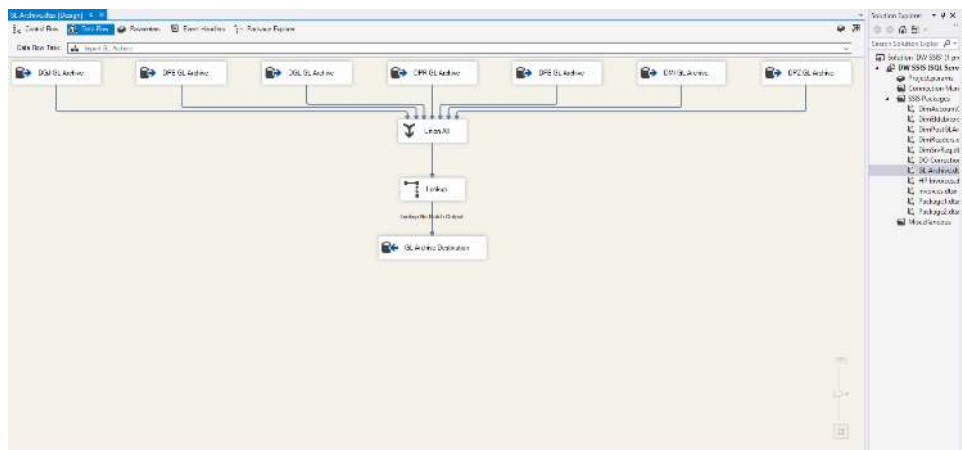


Figura 21. Data Flow i tabelës fakt 'GLArchive'

Tabela fakt ‘DOCorrections’

Pyetësi i përdorur për të mbushur këtë tabelë është:

WHERE YEAR(td.ReportRate) > 2015

exec updateFactCorrectionDebitor – për plotësimin e shënimeve nga
tabela dimension 'ElDebitors'

Paketa në SSIS për tabelën fakt 'DOCorrections'

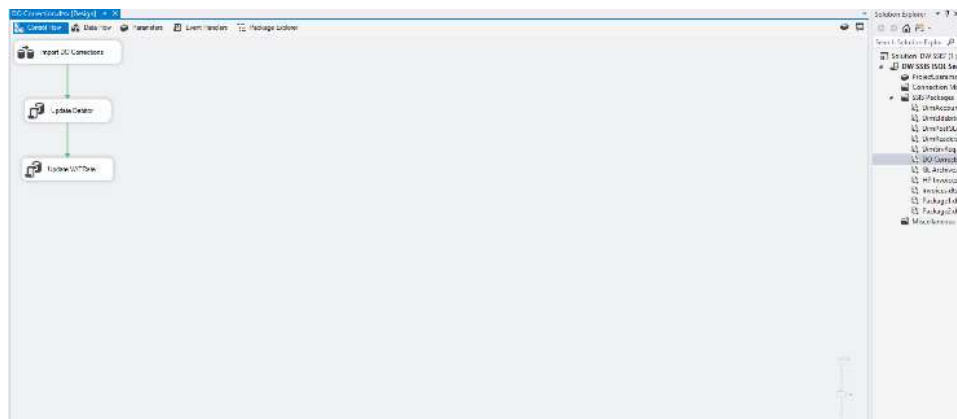


Figura 22. Control Flow i tabelēs fakt 'DOCorrections'

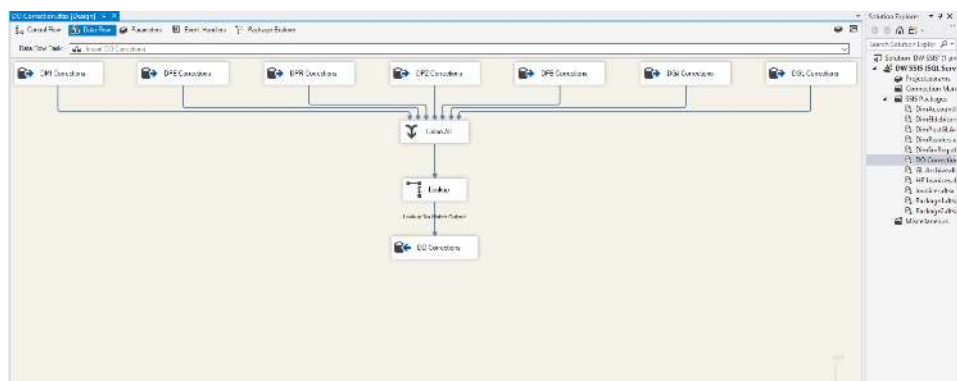


Figura 23. Data Flow i tabelēs fakt 'DOCorrections'

3.17 Konstruktimi i kubeve

Pas procesit të bartjes së të dhënave nga bazat e të dhënave transaksionale përmes ETL, do të vazhdojmë me krijimin e kubeve dhe përgatitjen e të dhënave për fazën e raportimit.

3.17.1 Krijimi i DSV

DSV është abstraksion i një burimi relacional të të dhënave që bëhet bazë e kubeve dhe dimensioneve në projektin multidimensional. Qëllimi i DSV-së është të mundësojë kontrollin mbi strukturën e të dhënave që përdoret në projekt.

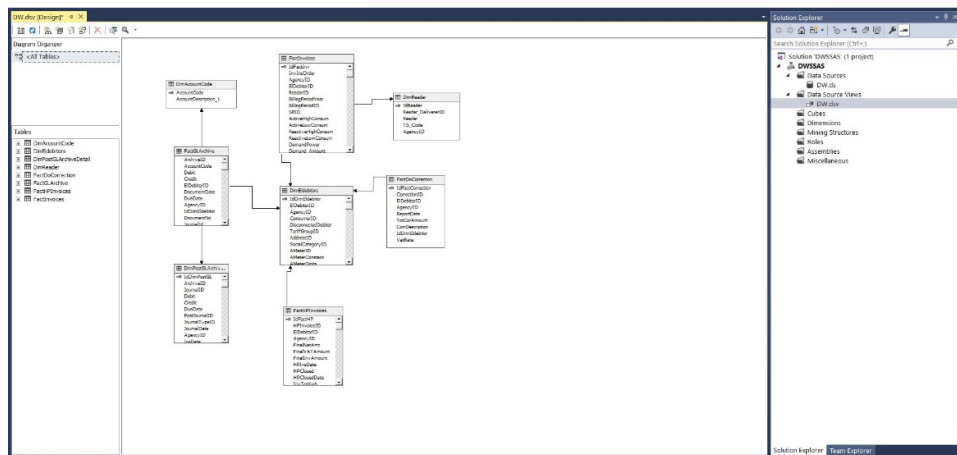


Figura 24. Krijimi i DSV

3.17.2 Krijimi i kubit

Pas definimit të DSV, hapi tjetër është definimi i kubit dhe dimensioneve të cilat do të përdoren nga kubi. Në dimensione do t'i bartim të gjitha tabelat dimension të cilat i kemi krijuar përmes SSIS. Kubi do të përmbajë tabelat fakt të krijuara në SSIS.

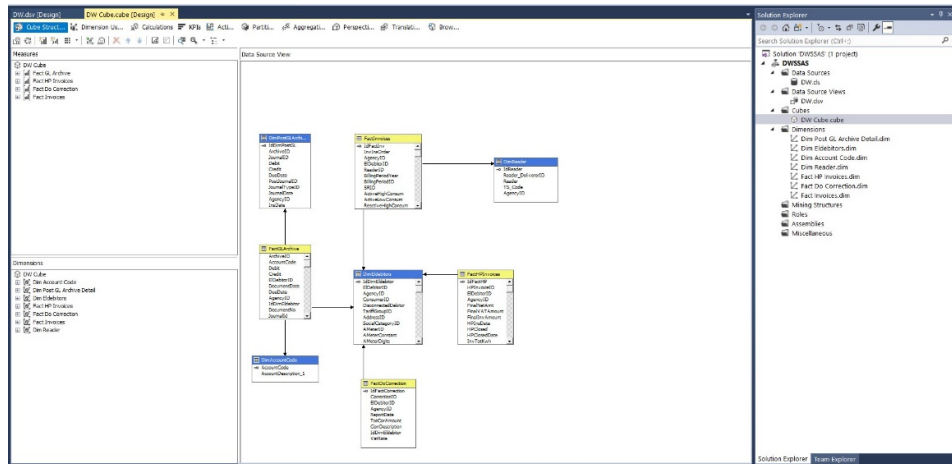


Figura 25. Krijimi i kubit

Dimenzionet mund të përdoren për një apo më shumë kube. Në figurën e mëposhtme paraqitet lidhja e dimenzioneve me grupet matëse.

DW Cube.cube [Design]				
Measure Groups				
Dimensions	[u] Fact GL Archive	[u] Fact HP Invoices	[u] Fact Do Correction	[u] Fact Invoices
[u] Dim Account Code	Account Code			
[u] Dim Post GL Archive D...	Id Dim Post GL			
[u] Dim Eldebitors	Id Dim Eldebitor	Id Dim Eldebitor	Id Dim Eldebitor	Id Dim Eldebitor
[u] Fact HP Invoices		Id Fact HP		
[u] Fact Do Correction			Id Fact Correction	
[u] Fact Invoices				Id Fact Inv
[u] Dim Reader				Id Reader

Figura 26. Lidhja e dimenzioneve me grupet matëse

3.17.3 Definimi i kalkulimeve

Anëtarët e kalkuluar janë anëtarë të dimenzioneve apo grupeve matëse që janë të definuar bazuar në kombinimin e të dhënave të kubit, operacioneve aritmetike, numrave dhe funksioneve.

Një shembull i një kalkulimi ndërmjet dy elementeve të kubit është paraqitur në figurën e mëposhtme.

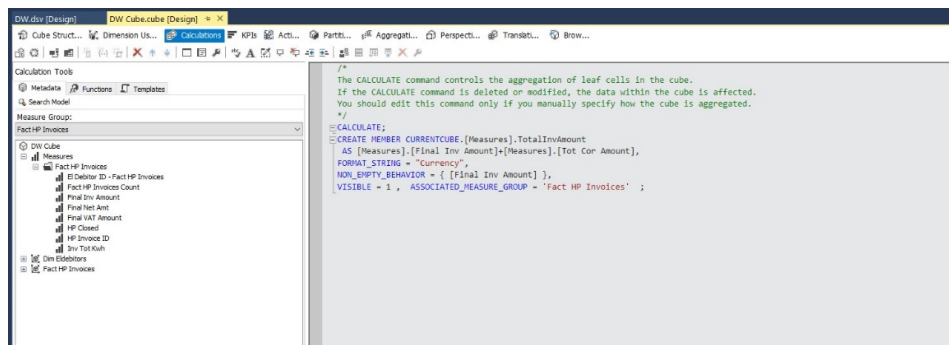


Figura 27. Kalkulimi

3.17.4 Definimi i KPI

Indikatorët kyç të performancës të njohur si KPI standardizojnë paraqitjen e rezultateve. KPI-të mund të shfaqen në raporte, portale, panele, API, etj. Një KPI mund të krijohet nga një apo më shumë grupe matëse.

Së pari caktohen shprehjet e vlerave apo kalkulimet, pastaj shprehja e targetit që synohet, statusi i cili paraqitet grafikisht për të parë më lehtë statusin e shprehjes së vlerave me vlerën e synuar, dhe në fund trendi apo tendenca e cila tregon se vlerat e kalkuluara janë duke u përmirësuar apo duke shkuar më keq në relacion me vlerat e synuara.

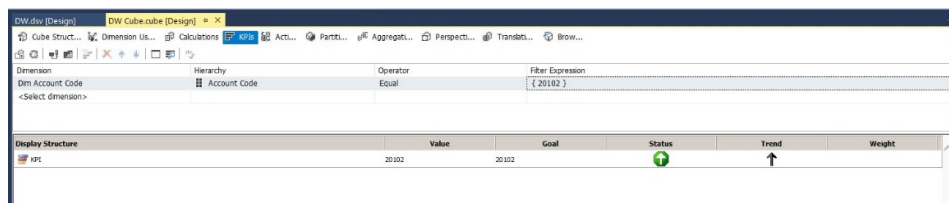


Figura 28. Krijimi i KPI

3.17.5 Definimi i aksioneve

Për të rritur performancën e pyetësorëve duhet të definohen aksionet të cilat mundësojnë filtrimin e të dhënave duke caktuar numrin total të rreshtave që kthehen si rezultat. Shembulli i krijimit të një aksioni 'drillthrough' është paraqitur në figurën e mëposhtme.

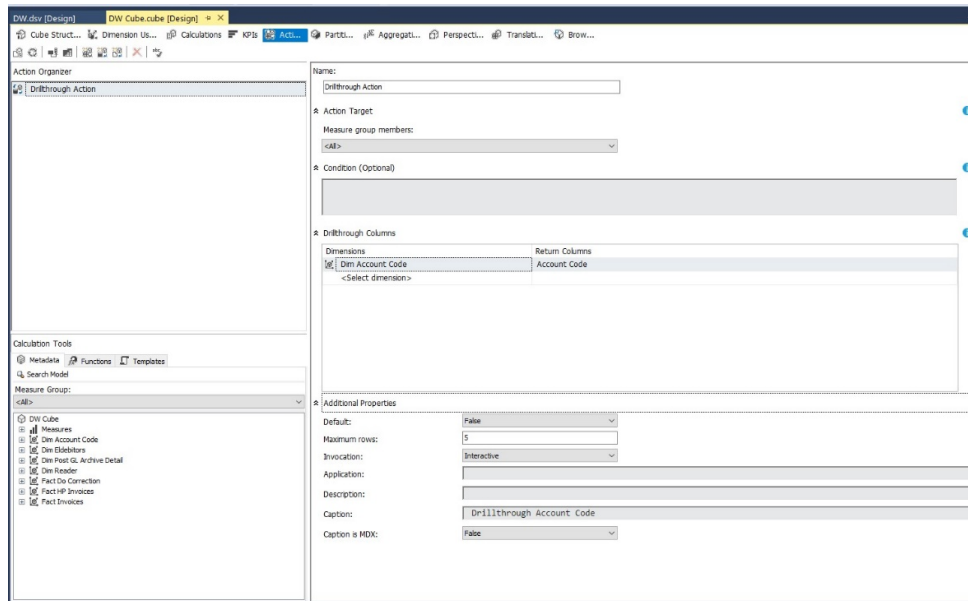


Figura 29. Krijimi i aksionit ‘drillthrough’

3.18 Raportet

Tani që ekziston prototipi i data warehouse, pyetja është se çfarë mund të bëjmë me të dhënat për t’i paraqitur përparësitë e data warehouse.

Raportet të cilat i kemi përzgjedhur për t’i gjeneruar në këtë rast studimi, do t’i paraqesim edhe përmes veglës SQL Server Reporting Services (SSRS) me të cilën do të ekzekutojmë pyetësorët për shfaqjen e rezultateve të fituara nga procedurat e ekzekutuara në datawarehouse.

3.18.1 Faturimi i kWh dhe numri i konsumatorëve

Pyetësoni i përdorur për paraqitjen e këtij raporti është:

```
SELECT fb.District, fb.Trafo,
fb.[4/01] AS TariffGroupID, fb.[1] AS NrCons, fb.[2] AS [A+], fb.[3] AS [A-],
fb.[4/02] AS TariffGroupID, fb.[4] AS NrCons, fb.[5] AS [A+], fb.[6] AS [A-],
fb.[6/6] AS TariffGroupID, fb.[7] AS NrCons, fb.[8] AS [A+], fb.[9] AS [A-],
```


fb.[7/01] AS TariffGroupID,fb.[10] AS NrCons,fb.[11] AS [A+],fb.[12] AS [A-],

fb.[7/02] AS TariffGroupID,fb.[13] AS NrCons,fb.[14] AS [A+],fb.[15] AS [A-],

fb.[8/01] AS TariffGroupID,fb.[16] AS NrCons,fb.[17] AS [A+],fb.[18] AS [A-],

fb.[1/2] AS TariffGroupID,fb.[19] AS NrCons,fb.[20] AS [A+],fb.[21] AS [A-],

fb.[1/3] AS TariffGroupID,fb.[22] AS NrCons,fb.[23] AS [A+],fb.[24] AS [A-],

fb.[9/1] AS TariffGroupID,fb.[25] AS NrCons,fb.[26] AS [A+],fb.[27] AS [A-],

fb.[9/2] AS TariffGroupID,fb.[28] AS NrCons,fb.[29] AS [A+],fb.[30] AS [A-],

fb.[9/3] AS TariffGroupID,fb.[31] AS NrCons,fb.[32] AS [A+],fb.[33] AS [A-]

FROM FaturimiHQ fb

ORDER BY District,Trafo

Dhe forma në të cilën paraqitet në SSRS është:

Faturimi HQ																							
Sublet	Tudo	TS 401	TS 402	TS 403	TS 404	TS 405	TS 406	TS 407	TS 408	TS 409	TS 410	TS 411											
TS1	810000	401	38.0000	3375.0000	0.0000	4002	700.0000	1781780.0000	163377.0000	66	65.0000	201001.2000	74792.0000	701	21.0000	951.0000	0.0000	TS2	1790.0000	533814.0000	215489.0000	53.0000	1989
TS2	180140	401	55.0000	8554.0000	0.0000	4002	1800.0000	232024.0000	161800.0000	66	4.0000	13048.0000	8153.0000	701	3.0000	19500.0000	0.0000	TS2	90.0000	35306.0000	19851.0000	6.0000	1200
TS3	400000	401	191.0000	33371.0000	0.0000	4002	9400.0000	2420000.0000	1600000.0000	66	44.0000	40000.0000	100000.0000	701	70.0000	42000.0000	0.0000	TS2	1770.0000	573878.0000	207000.0000	54.0000	6300
TS4	400400	401	83.0000	22988.0000	0.0000	4002	5100.0000	1401100.0000	841244.0000	66	11.0000	110000.0000	41001.0000	701	60.0000	25000.0000	0.0000	TS2	1760.0000	754274.0000	202000.0000	25.0000	2500
TS5	400410	401	100.0000	10741.0000	0.0000	4002	5770.0000	1548000.0000	1004000.0000	66	60.0000	72200.0000	245771.0000	701	30.0000	1710.0000	0.0000	TS2	1000.0000	311120.0000	107770.0000	40.0000	5500
TS6	400420	401	110.0000	12004.0000	0.0000	4002	8194.0000	2171000.0000	1200000.0000	66	40.0000	210000.0000	107004.0000	701	30.0000	43000.0000	0.0000	TS2	1000.0000	343863.0000	118001.0000	32.0000	3504
TS7	400430	401	114.0000	21821.0000	0.0000	4002	8770.0000	1400000.0000	117014.0000	66	10.0000	10000.0000	20000.0000	701	20.0000	2120.0000	0.0000	TS2	800.0000	243200.0000	100000.0000	17.0000	1100
TS8	400440	401	154.0000	12419.0000	0.0000	4002	3800.0000	730000.0000	104514.0000	66	12.0000	41700.0000	7000.0000	701	6.0000	1014.0000	0.0000	TS2	440.0000	140071.0000	100001.0000	34.0000	3772
TS9	400450	401																TS2					
TS10	400460	401																TS2					
TS11	400470	401																TS2					
TS12	400480	401																TS2					
TS13	400490	401																TS2					
TS14	400500	401																TS2					
TS15	400510	401																TS2					
TS16	400520	401																TS2					
TS17	400530	401																TS2					
TS18	400540	401																TS2					
TS19	400550	401																TS2					
TS20	400560	401																TS2					
TS21	400570	401																TS2					
TS22	400580	401																TS2					
TS23	400590	401																TS2					
TS24	400600	401																TS2					
TS25	400610	401																TS2					
TS26	400620	401																TS2					
TS27	400630	401																TS2					
TS28	400640	401																TS2					
TS29	400650	401																TS2					
TS30	400660	401																TS2					
TS31	400670	401																TS2					
TS32	400680	401																TS2					
TS33	400690	401																TS2					
TS34	400700	401																TS2					
TS35	400710	401																TS2					
TS36	400720	401																TS2					
TS37	400730	401																TS2					
TS38	400740	401																TS2					
TS39	400750	401																TS2					
TS40	400760	401																TS2					
TS41	400770	401																TS2					
TS42	400780	401																TS2					
TS43	400790	401																TS2					
TS44	400800	401																TS2					
TS45	400810	401																TS2					
TS46	400820	401																TS2					
TS47	400830	401																TS2					
TS48	400840	401																TS2					
TS49	400850	401																TS2					
TS50	400860	401																TS2					
TS51	400870	401																TS2					
TS52	400880	401																TS2					
TS53	400890	401																TS2					
TS54	400900	401																TS2					
TS55	400910	401																TS2					
TS56	400920	401																TS2					
TS57	400930	401																TS2					
TS58	400940	401																TS2					
TS59	400950	401																TS2					
TS60	400960	401																TS2					
TS61	400970	401																TS2					
TS62	400980	401																TS2					
TS63	400990	401																TS2					
TS64	401000	401																TS2					
TS65	401010	401																TS2					
TS66	401020	401																TS2					
TS67	401030	401																TS2					
TS68	401040	401																TS2					
TS69	401050	401																TS2					
TS70	401060	401																TS2					
TS71	401070	401																TS2					
TS72	401080	401																TS2					
TS73	401090	401																TS2					
TS74	401100	401																TS2					
TS75	401110	401																TS2					
TS76	401120	401																TS2					
TS77	401130	401																TS2					
TS78	401140	401																TS2					
TS79	401150	401																TS2					
TS80	401160	401																TS2					
TS81	401170	401																TS2					
TS82	401180	401																TS2					
TS83	401190	401																TS2					
TS84	401200	401																TS2					
TS85	401210	401																TS2					
TS86	401220	401																TS2					
TS87	401230	401																TS2					
TS88	401240	401																TS2					
TS89	401250	401																TS2					
TS90	401260	401	</																				

Agency ID	TS Code	A1 Consum	A2 Consum	R1 Consum	R2 Consum	Consumer No
DFE	10014011001	15496.00	16443.00	0.00	0.00	76
DFE	10014011002	10966.50	8299.10	0.00	10.92	38
DFE	10014011003	28044.00	22000.00	0.00	0.00	127
DFE	10014011004	15553.00	9834.00	0.00	0.00	59
DFE	10014011005	11517.00	5752.00	0.00	0.00	42
DFE	10014011006	27212.00	14357.00	0.00	0.00	116
DFE	10014011007	16745.00	6811.40	3470.34	2.69	43
DFE	10014011008	10493.00	9878.00	0.00	0.00	60
DFE	10014011009	20916.00	9504.00	0.00	0.00	99
DFE	10014011010	22690.00	8701.00	0.00	0.00	111

Figura 31. Raporti i faturimit TS në SSRS

3.18.3 Pagesat

Pyetësoi i përdorur për paraqitjen e këtij raporti është:

```

DECLARE @from AS DATETIME,@to AS DATETIME

SELECT @from = '20180401',@to = '20180410 23:59'

SELECT d.AgencyID, d.TariffGroupID, SUM(tpf.Credit) AS TotalCollection
FROM FactGLArchive tpf INNER JOIN DimEldebtors d
ON d.AgencyID=tpf.AgencyID and d.ElDebitorID = tpf.ElDebitorID
WHERE tpf.DueDate BETWEEN @from AND @to
AND d.TariffGroupID<>'3' AND tpf.AccountCode in (20102,201022)

GROUP BY d.AgencyID,d.TariffGroupID

ORDER BY d.AgencyID,d.TariffGroupID

```

Dhe forma në të cilën paraqitet në SSRS është:

Agency ID	Tariff Group ID	Total Col
DFE	1/2	1089.0500
DFE	1/3	84457.5300
DFE	4/01	3069.9700
DFE	4/02	499445.3900
DFE	6/6	34441.7000
DFE	7/01	972.2600
DFE	7/02	100432.0500
DFE	8/01	20.0000
DGJ	1/3	15833.8400
DGJ	4/01	250.5000
DGJ	4/02	429550.7610
DGJ	6/6	24393.9000

Figura 32. Raporti i pagesave në SSRS

3.18.4 Raporti i faturave zero

Pyetësori i përdorur për paraqitjen e këtij raporti është:

```
SELECT AgencyID, SUM(TotalZeros) AS TotalZeros, SUM(NrPayments) AS
NrPayments, SUM(TotalPaid) AS TotalPaid, SUM(TotalBA) AS TotalBA,
SUM(TotalDebt) AS TotalDebt, SUM(KedsDebt) AS KedsDebt, SUM(No_E)
AS NoE, SUM(No_M) AS NoM, SUM(TotalBills) AS TotalBills

FROM ReportZeros

GROUP BY AgencyID

ORDER BY 1
```

Dhe forma në të cilën paraqitet në SSRS është:

District	Total Zeros	Nr Payments	Total P	Total Ba	Total D	Keds D	No E	No M	Total B
DFE	138510	35898.00	1325367.75	1290490.92	8543325.50	2402202.56	1526	181866	183392
DGJ	96725	25969.00	1021508.69	974526.11	4303434.06	1258198.94	786	125649	126435
DGL	114308	30429.00	1090103.15	1049529.68	5974790.72	1637020.56	1191	148627	149818
DMI	97877	27053.00	795579.30	823642.18	7022888.00	2023354.81	1426	128482	129908
DPE	151613	29238.00	1296072.53	1337147.10	7682283.24	2866791.28	1474	189761	191235
DPR	252371	72484.00	2692048.26	2722059.15	15010934.64	5264533.32	6050	330919	336969
DPZ	178371	42691.00	1690066.07	1588464.99	7820582.07	2534853.33	1673	226684	228357

Figura 33. Raporti i faturave zero në SSRS

Kapitulli IV

4 Vlerësimi i performancës

4.1 Paraqitja e përgjithshme

Qëllimi i këtij testimi është të mirret një shembull konkret për të paraqitur përmirësimin në kohë të pyetësorit, duke e krahasuar bazën e të dhënave transaksionale me star skemën e denormalizuar në data warehouse.

Me qëllim që data warehouse të bëhet një opsion i mundshëm dhe për ta bërë atë një përpjekje të vlefshme, ajo është e rëndësishme për të treguar se do të ketë një përmirësim të madh në kohën e pyetësorëve.

Kjo përpjekje përfshinë zgjidhjen e një problemi që ekzistonte ku zgjidhja ishte krijimi i një skeme yll në data warehouse. Në një bazë të madhe të të dhënave, koha e gjenerimit të raportit mvaret nga kompleksiteti i lidhjes së tabelave. Qëllimi ishte të kemi një tabelë fakt që korrespondon me rekordet që janë në bazë të të dhënave. Tabela fakt lidhet me tabelat dimension të cilat përmbajnë atributet që i paraqet një view.

Pritshmëria nga ky test është që të dhënat të cilat kthehen nga data warehouse marrin më pak kohë se në bazën e të dhënave transaksionale. Arsyeja që do të kemi dallim në kohë është nga mënyra e organizimit në data warehouse.

4.2 Plani testues

Në mënyrë që t'i shohim ndryshimet ndërmjet dy sistemeve, kemi bërë teste për t'i krahasuar pyetësorët. Vegla e përdorur për testim është T-SQL. Pyetësorët e njëjtë do t'i ekzekutojmë në të dy sistemet dhe rezultatet e fituara do t'i shfaqim më poshtë.

4.2.1 Faturimi i kWh dhe numri i konsumatorëve

Raporti për faturimin kWh dhe numrin e konsumatorëve të grupuara sipas distriktit, kodit të TS dhe grupit tarifyor. Nga ekzekutimi i pyetësorëve të njëjtë në të dy sistemet shohim se në data warehouse marrim rezultatin për kohë më të shkurtër. Në figurat e mëposhtme janë rezultatet e fituara.

```
SQLQuery32.sql - 10...RTS (nertilak (58))* x SQLQuery29.sql - 10...DW (nertilak (60))*
142 SET f.[9/2] = '9/2',
143 f.[28] = c.nrcons,
144 f.[29] = c.[a+],
145 f.[30] = c.[a-]
146 FROM nk.faturimihq f INNER JOIN #t c ON f.Trafo=c.trafo AND f.District=c.AgencyId
147 WHERE c.tariffgroupid='9/2'
148
149 UPDATE f
150 SET f.[9/3] = '9/3',
151 f.[31] = c.nrcons,
152 f.[32] = c.[a+],
153 f.[33] = c.[a-]
154 FROM nk.faturimihq f INNER JOIN #t c ON f.Trafo=c.trafo AND f.District=c.AgencyId
155 WHERE c.tariffgroupid='9/3'
156
157
158 ----Select
159
```

100 %

Messages

(523 row(s) affected)

(99 row(s) affected)

(68 row(s) affected)

(73 row(s) affected)

(80 row(s) affected)

(69 row(s) affected)

(75 row(s) affected)

100 %

Query executed successfully. 10.150.1.112/sql2005dev (9... nertilak (58) REPORTS 00:00:09 rows

Figura 34. Faturimi kWh në bazën e të dhënave transaksionale

```

SQLQuery32.sql - 10...RTS (nertilak (58))*  SQLQuery29.sql - 10...DW (nertilak (60))*
1 CREATE TABLE #t
2 (
3     Id INT IDENTITY,
4     AgencyId NVARCHAR(10),
5     Trafo NVARCHAR(11),
6     TariffGroupID NVARCHAR(10),
7     NrCons INT,
8     [a+] DECIMAL(15,4),
9     [a-] DECIMAL(15,4)
10 )
11 --Insert
12 ;WITH cte AS
13 (
14     SELECT tch.AgencyID, LEFT(d.TS_Code,6) AS Trafo,tch.TariffGroupID,
15     COUNT(distinct tch.ElDebitorID) AS NrCons,SUM(tch.ActiveHighConsum) AS [A+],SUM(tch.ActiveLowConsum) AS [A-]
16     FROM FactInvoices tch
17     inner join DimReader d on tch.IdReader=d.IdReader
18 )
19 INSERT INTO #t
20 SELECT * FROM cte
21 
```

Messages

(523 row(s) affected)

(99 row(s) affected)

(68 row(s) affected)

(73 row(s) affected)

(80 row(s) affected)

(69 row(s) affected)

(75 row(s) affected)

Query executed successfully. | 10.150.1.161 (11.0 SP1) | nertilak (60) | DW | 00:00:03 | 0 rows

Figura 35. Faturimi kWh në datawarehouse

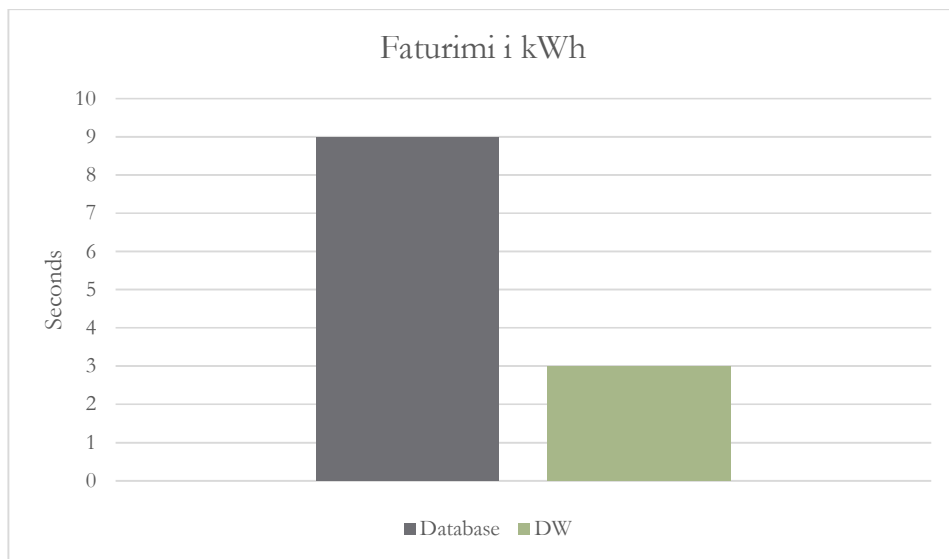


Figura 36. Rezultatet grafike të faturimit kWh

4.2.2 Faturimi i TS

Raporti për faturimin e energjisë aktive dhe energjisë reaktive të grupuara sipas distriktit dhe kodit të TS. Edhe në këtë rast shohim se pyetësi

ekzekutohet më shpejtë në data warehouse. Më poshtë janë figurat me kohën e ekzekutimit në të dy sistemet.

SQLQuery44.sql - 10...RTS (nertlak (61))

```

1 ;with cte as (
2   select Reader_DelivererID,Reader,BillingGroup,EmployeeID,TS_Code,TS_Code_11,'DFE' as AgencyID
3   from cisp_dfe.dbo.tblccreader
4   union all
5   select Reader_DelivererID,Reader,BillingGroup,EmployeeID,TS_Code,TS_Code_11,'DGJ' as AgencyID
6   from cisp_dgj.dbo.tblccreader
7   union all
8   select Reader_DelivererID,Reader,BillingGroup,EmployeeID,TS_Code,TS_Code_11,'DGL' as AgencyID
9   from cisp_dgl.dbo.tblccreader
10  union all
11  select Reader_DelivererID,Reader,BillingGroup,EmployeeID,TS_Code,TS_Code_11,'DMI' as AgencyID
12  from cisp_dmi.dbo.tblccreader
13  union all
14  select Reader_DelivererID,Reader,BillingGroup,EmployeeID,TS_Code,TS_Code_11,'DPE' as AgencyID
15  from cisp_dpe.dbo.tblccreader
16  union all
17  select Reader_DelivererID,Reader,BillingGroup,EmployeeID,TS_Code,TS_Code_11,'DPR' as AgencyID

```

AgencyID	TS_Code	A1_Consum	A2_Consum	R1_Consum	R2_Consum	ConsumerNo
DFE	10014011001	15496.00	16443.00	0.00	0.00	76
DFE	10014011002	10966.50	8299.10	0.00	10.92	38
DFE	10014011003	28044.00	22000.00	0.00	0.00	127
DFE	10014011004	15553.00	9834.00	0.00	0.00	59
DFE	10014011005	11517.00	5752.00	0.00	0.00	42
DFE	10014011006	27212.00	14357.00	0.00	0.00	116
DFE	10014011007	16745.00	6811.40	3470.34	2.69	43
DFE	10014011008	10493.00	9878.00	0.00	0.00	60
DFE	10014011009	20916.00	9504.00	0.00	0.00	99
DFE	10014011010	22690.00	8701.00	0.00	0.00	111
DFE	10014011011	8181.00	3106.00	0.00	0.00	59

Query executed successfully. 10.150.1.112(sql2005dev (9... nertlak (61) REPORT 00:00:09 639 rows

Figura 37. Faturimi TS në bazën e të dhënave transaksionale

SQLQuery44.sql - 10...RTS (nertlak (61))

```

1 select Reader_DelivererID,Reader,TS_Code,AgencyID
2 into #readers
3 from DimReader
4
5
6
7 DECLARE @billingperiodid INT
8 DECLARE @billingperiodyear INT
9 SET @billingperiodid=4
10 SET @billingperiodyear=2018
11
12 SELECT i.AgencyID,i.ReaderID
13 ,SUM(i.ActiveHighConsum ) AS ActiveHighConsum
14 ,SUM(i.ActiveLowConsum ) AS ActiveLowConsum
15 ,SUM(i.ReactiveHighConsum) AS ReactiveHighConsum
16 ,SUM(i.ReactiveLowConsum ) AS ReactiveLowConsum
17 ,COUNT(DISTINCT i.ElDebitorID) AS ConsumerNo

```

AgencyID	TS_Code	A1_Consum	A2_Consum	R1_Consum	R2_Consum	ConsumerNo
DFE	10014011001	15496.00	16443.00	0.00	0.00	76
DFE	10014011002	10966.50	8299.10	0.00	10.92	38
DFE	10014011003	28044.00	22000.00	0.00	0.00	127
DFE	10014011004	15553.00	9834.00	0.00	0.00	59
DFE	10014011005	11517.00	5752.00	0.00	0.00	42
DFE	10014011006	27212.00	14357.00	0.00	0.00	116
DFE	10014011007	16745.00	6811.40	3470.34	2.69	43
DFE	10014011008	10493.00	9878.00	0.00	0.00	60
DFE	10014011009	20916.00	9504.00	0.00	0.00	99
DFE	10014011010	22690.00	8701.00	0.00	0.00	111
DFE	10014011011	8181.00	3106.00	0.00	0.00	59

Query executed successfully. 10.150.1.161 (11.0 SP1) nertlak (56) DW 00:00:06 639 rows

Figura 38. Faturimi i TS në data warehouse

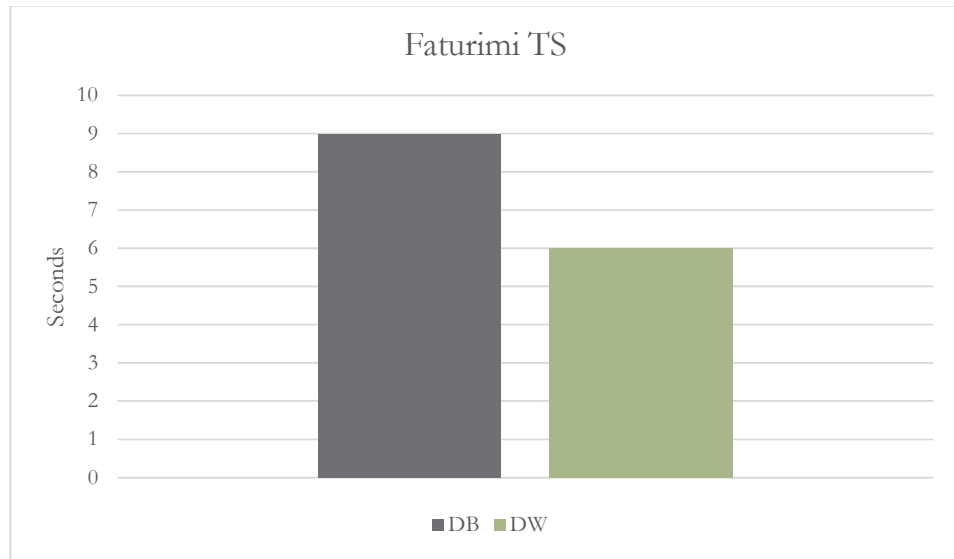


Figura 39. Diagrami i krahasimit të faturimit të TS

4.2.3 Pagesat

Raporti për pagesat e konsumatorëve të grupuara sipas distriktit dhe grupit tarifor në bazë të datave të caktuara. Edhe në këtë krahasim shohim se dukshëm më shpejtë kthehen rezultatet nga pyetësi i ekzekutuar në data warehouse.

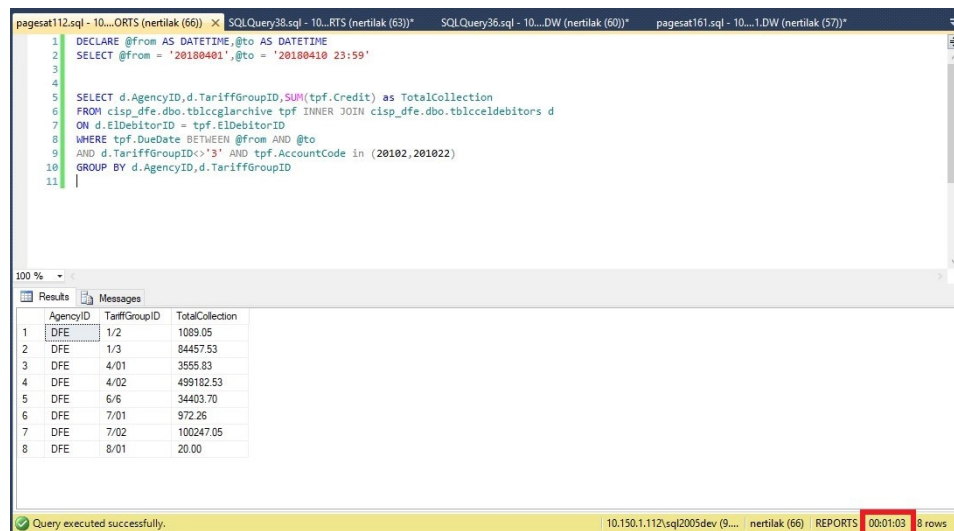


Figura 40. Pagesat në bazën e të dhënave transaksionale

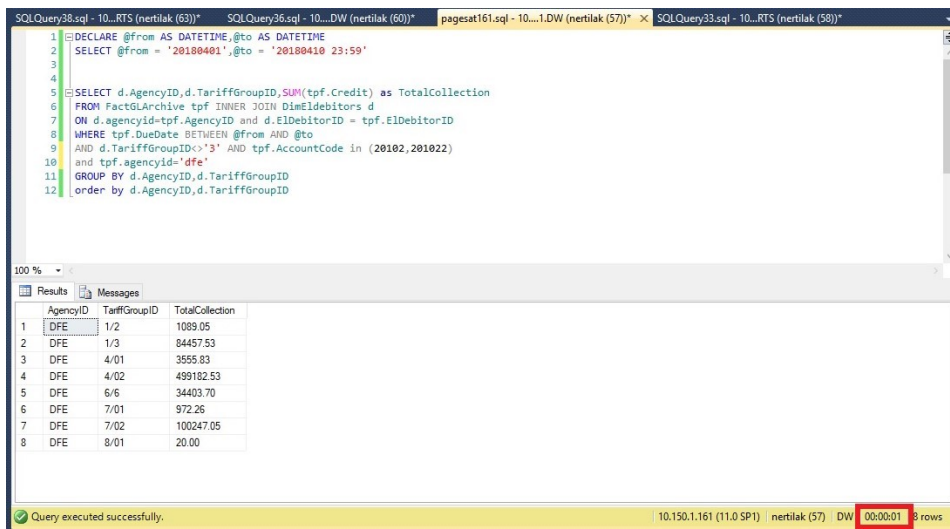


Figura 41. Pagesat në data warehouse

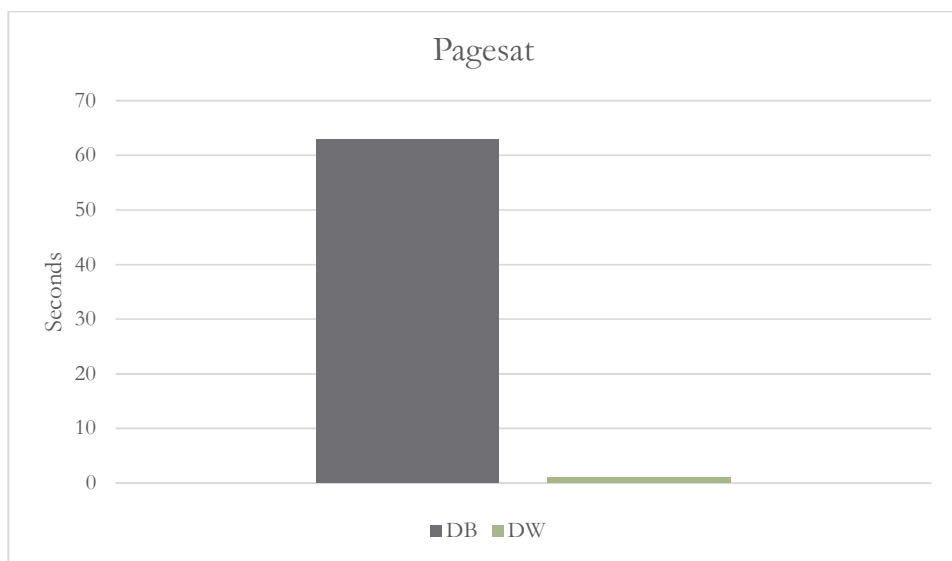


Figura 42. Diagrami i rezultateve të pagesave

4.2.4 Raporti i faturave zero

Raporti i faturimit të konsumatorëve pa shpenzim. Ky raport bazohet në shumë kalkulime dhe ekzekutohen disa procedura për rezultatin final. Procedurat e njejta janë ekzekutuar në të dy sistemet dhe shohim se në data warehouse e fitojmë rezultatin në kohë më të shkurtë.

```
SQLQuery1.sql - 10...ORTS (nertlak (54)) *  
1 USE [REPORTS]  
2 GO  
3 /***** Object: StoredProcedure [NK].[reportzerose] Script Date: 31.05.2018 19:41:12 PM *****/  
4 SET ANSI_NULLS ON  
5 GO  
6 SET QUOTED_IDENTIFIER ON  
7 GO  
8 ALTER procedure [NK].[reportzerose]  
9 as  
10 begin  
11 exec nk.ReportZerosAllIDFE  
12 exec nk.ReportZerosAllIDGJ  
13 exec nk.ReportZerosAllIDGL  
14 exec nk.ReportZerosAllIMI  
15 exec nk.ReportZerosAllIPE  
16 exec nk.ReportZerosAllPr  
17 exec nk.ReportZerosAllIPZ  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012  
2013  
2014  
2015  
2016  
2017  
2018  
2019  
2020  
2021  
2022  
2023  
2024  
2025  
2026  
2027  
2028  
2029  
2030  
2031  
2032  
2033  
2034  
2035  
2036  
2037  
2038  
2039  
2040  
2041  
2042  
2043  
2044  
2045  
2046  
2047  
2048  
2049  
2050  
2051  
2052  
2053  
2054  
2055  
2056  
2057  
2058  
2059  
2060  
2061  
2062  
2063  
2064  
2065  
2066  
2067  
2068  
2069  
2070  
2071  
2072  
2073  
2074  
2075  
2076  
2077  
2078  
2079  
2080  
2081  
2082  
2083  
2084  
2085  
2086  
2087  
2088  
2089  
2090  
2091  
2092  
2093  
2094  
2095  
2096  
2097  
2098  
2099  
2100  
2101  
2102  
2103  
2104  
2105  
2106  
2107  
2108  
2109  
2110  
2111  
2112  
2113  
2114  
2115  
2116  
2117  
2118  
2119  
2120  
2121  
2122  
2123  
2124  
2125  
2126  
2127  
2128  
2129  
2130  
2131  
2132  
2133  
2134  
2135  
2136  
2137  
2138  
2139  
2140  
2141  
2142  
2143  
2144  
2145  
2146  
2147  
2148  
2149  
2150  
2151  
2152  
2153  
2154  
2155  
2156  
2157  
2158  
2159  
2160  
2161  
2162  
2163  
2164  
2165  
2166  
2167  
2168  
2169  
2170  
2171  
2172  
2173  
2174  
2175  
2176  
2177  
2178  
2179  
2180  
2181  
2182  
2183  
2184  
2185  
2186  
2187  
2188  
2189  
2190  
2191  
2192  
2193  
2194  
2195  
2196  
2197  
2198  
2199  
2200  
2201  
2202  
2203  
2204
```

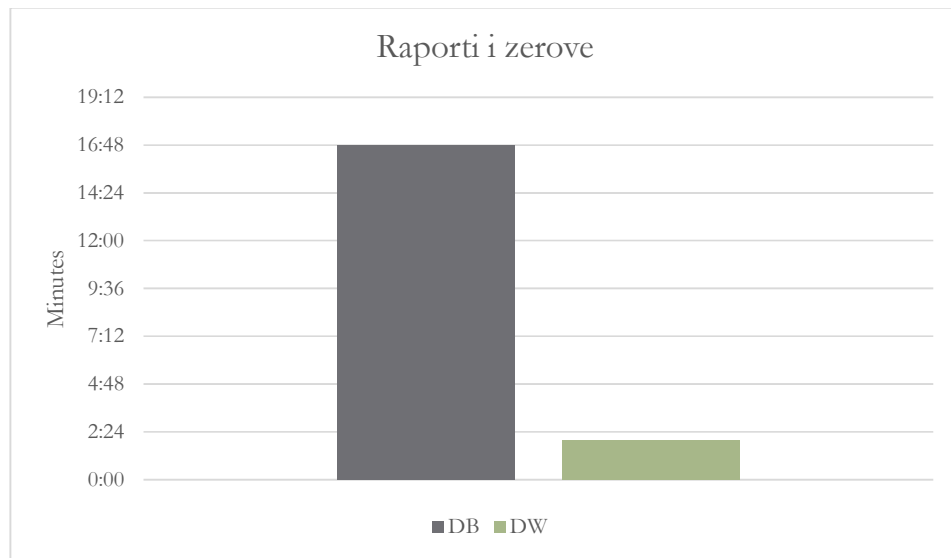


Figura 45. Diagrami i rezultateve të zerove

Kapitulli V

5 Konkluzioni

Kjo tezë paraqet idenë e implementimit të data warehouse në kompaninë KEDS, në mënyrë që të integrohen më mirë sistemet për të përfutur raportim më të thjeshtë dhe më të përmirësuar. Për të implementuar data warehouse në kompani është e domosdoshme që të ketë një angazhim të madh të kompanisë të udhëhequr nga një ekzekutiv i cili beson në projekt dhe është përgjegjës për mbajtjen e ekipit të projektit të motivuar. Departamenti i IT-së ekzekuton projektin, instalon data warehouse në kompani dhe integron sistemet ekzistuese.

Implementimi i data warehouse në kompani është më i efektshëm me përdorimin e teknikave më të reja të modelimit të të dhënave dhe praktikave të studiuara si më të mirat ndër vite, që janë baza e data warehouse në të gjithë botën. Duke i organizuar të dhënat në përmbledhje të skemës ‘snowflake’ në formë të tabelave fakt dhe dimenzion për organizim efektiv për nevoja të kompanisë, përmirësohet raportimi.

Benefitet kryesore nga implementimi i data warehouse janë listuar më poshtë:

- Në një tabelë dimenzion mund të bëhen thirrje nga shumë tabela fakt.
- Përdoruesit do të kursejnë kohë pasi që të dhënat historike përfitohen më shpejtë.
- Sistemi mund të përgjigjet në shumë pyetje njëkohësisht. Përdoruesit mund të gjenerojnë raporte të reja duke përdorur pyetësorët ekzistues apo duke i krijuar pyetësorët e tyre.
- Administratori i bazës së të dhënave lehtë mund të përditësojë kushtet e reja në tabelat e data warehouse.

Duke dizajnuar dhe implementuar një shkallë të vogël, një mini data warehouse në kompani, ishte e thjeshtë të shohim se si skema ‘snowflake’ me anë të veglave ETL mbledh shënime nga shumë burime. Me shembuj praktik pamë se si ekzekutohen pyetësorët e njëjtë me të dhëna të njëjta por në sisteme të ndryshme. Gjatë krahasimit të rezultateve u pa qartë se pyetësorët ekzekutohen më shpejtë në data warehouse se në baza të të dhënave transaksionale.

Shembujt praktik dhe studimi i arkitekturës dhe raportimit në data warehouse, shfaqin përparësitë e implementimit të data warehouse në kompani për një eksperiencë më të mirë për përdoruesit të cilët bëjnë pyetësorët dhe raportet.

Me një fjalë, implementimi i data warehouse në kompani të mëdha është më se i nevojshëm për planifikim dhe vendimmarrje më efikase.

Bibliografia

- [1] R. Kimball, L. Reeves, M. Ross dhe W. Thornthwaite, The Data Warehouse Lifecycle Toolkit, 1998.
- [2] P. O'Donnell, «Monash University,» 2004. [Në linjë]. Available: <http://www.sims.monash.edu.au/subjects/ims5024/ims5024/lectures/week10slides5024.pdf>.
- [3] P. Lane dhe V. Schupmann, «Oracle9i Data Warehousing Guide,» Oracle Pres, 2002.
- [4] T. Moss dhe S. Adelman, Data Warehouse Project Management, Addison-Wesley Information Technology Series, 2000.
- [5] H. Research, «Selecting a Data Warehouse Solution for a Two-Year College: Best Practices,» Hanover Research, Washington DC, 2012.
- [6] D. Rogers, «Data Warehousing Architecture - Designing the Data Staging Area,» *www.databasejournal.com*, 22 June 2010.
- [7] M. Usman, S. Asghar dhe S. Fong, «Data mining and automatic OLAP schema generation,» në *Digital Information Management (ICDIM), 2010 Fifth International Conference on*, Thunder Bay, ON, 2010.
- [8] M. Whitehorn, K. Burns dhe E. N. Hanson, «Best Practices for Data Warehousing with SQL Server 2008 R2,» December 2010.
- [9] D. Schaffhauser, «Florida State U Transforms Reporting with Business Intelligence,» Campus Technology, Florida, 2010.
- [10] D. Sarka, M. Lah dhe G. Jerkič, Exam 70-463:Implementing a Data Warehouse with Microsoft® SQL Server® 2012 Training Kit, California: Microsoft Corporation, 2012.
- [11] Corporation, Microsoft, «www.microsoft.com,» 26 May 2011. [Në linjë]. Available: <http://msdn.microsoft.com/en-us/library/bb399731.aspx>.

- [12] «www.data-warehouses.net,» 20 October 2011. [Në linjë]. Available: <http://data-warehouses.net/architecture/staging.html>.
- [13] T. Chrysafis, «On-Line Analytical Processing,» në *Second International Student Spring Symposium on Contemporary Topics in IT(CTIT)*, Thessaloniki, 2003.
- [14] I. Ahmad dhe S. Azhar, «Data Warehousing In Construction: From Conception To Application,» në *First International Conference on Construction in the 21st Century (CITC2002) Challenges and Opportunities in Management and Technology*, Miami, Florida, 2002.
- [15] M. Guidry, «Lynda: Implementing a Data Warehouse with Microsoft SQL Server 2012».
- [16] M. K. Tucker, «Data Warehousing in the Modern World: A Case Study Revisited,» SUGI 27 Proceedings, Orlando.
- [17] G. Mehler, «Data Warehousing for the Enterprise,» SUGI 27 Proceedings, Orlando.