



Datavetenskap

Mikael Andersson

Conny Nykvist

Mikael Morell

Mikael Forsberg

**Optimering av SQL-frågor för analys i
QlikView**

1 Sammanfattat omdöme av examensarbetet

1.1 Titel

Titeln speglar examensarbetet på ett bra sätt och väcker säkerligen intresse för de som är intresserade av databaser.

1.2 Uppsatsens disposition

Dispositionen av uppsatsen är bra i det stora hela.

1.3 Begreppsapparat

Överlag så är "svåra" begrepp förklarade på ett bra sätt. De begrepp som däremot saknas och skulle vara i behov av en förklaring är datatvätt, aggression, TKPROF, cursor och ODBC. Vid några tillfällen så saknas förklaringen av ett begrepp vid dess första förekomst i uppsatsen och det blir därmed svårt att följa med i texten bra. Det skulle även behövas till exempel en bilaga med en sammanställning över begrepp.

1.4 Argumentering och slutsatsdragning

Slutsatsdragningen från de argument som presenteras i avsnitt 7.1 känns bra, dock så kan man ifrågasätta argumenteringen till viss del. Problemet med att QlikView får för långa svarstider har inte, enligt vår syn, tagits upp i uppsatsen tidigare. Utförligare information om just det problemet vore önskvärt.

1.5 Sammanfattningen

Sammanfattningen är lättöverskådlig och belyser de viktiga delarna i uppsatsen, såsom syfte och resultat.

1.6 Språkbehandling

Språkbehandlingen är god, det finns dock vissa stavfel och så vidare i uppsatsen. Samt är någon enstaka mening lite rörig och svårförståelig. En tendens i de tidigare delarna av uppsatsen är att kommatecken används onödigt mycket, konsekvensen av detta är att vissa meningar blir väldigt svåra att förstå.

1.7 Referat och källförteckning

Hittar inte information om ETL (Extract Transform Load) som referens 11 skulle referera till.

I övrigt verkar de referenser som vi har gått igenom att stämma.

2 Genomgång av uppsatsen kapitelvis

2.1 Kapitel 1 Inledning

Tar upp uppsatsens innehåll kortfattat men ändå tillräckligt utförligt på ett bra sätt. Avsnittet om uppsatsens disposition var ett bra och lätt sätt för att få en snabb överblick över de i uppsatsen innefattande kapitlen och deras innehåll. Det som är lite oklart är exakt vad ”förväntningar och mål” innefattar (Kapitel 1, rad 14).

2.2 Kapitel 2 Teori

Går igenom de teoretiska delar som behövs för en bra förståelse av senare kapitel i uppsatsen.

Avsnitt 2.1 Datalager, sidan 4 rad 2-4 innehåller en mening som är lite rörig: ”Generellt sett kan ett datalager beskrivas som en samling av beslutsfattande teknologier, riktat att möjliggöra, för verkställande ledare, att fatta fördelaktigare och snabbare beslut.” Denna mening skulle behöva en omskrivning för att tydliggöra vad som menas. Det är även två stycken ord som behöver en närmare förklaring alternativt omskrivning, dessa ord är perception och nyckelbitar.

Avsnitt 2.1.1 Definition sidan 4 stycke 1 rad 1 innehåller en referens som är lite otydlig, är det en referens till författaren, till en bok om författaren, en bok skriven av författaren etc.

detta skulle behöva förtydligas så man inte behöver gissa med hjälp av den refererade bokens titel.

Sidan 4 stycke 3 punkt 2 rad 3 innehåller begreppet datatvätt men en förklaring av begreppet tas inte upp i uppsatsen och detta skulle behövas.

Avsnitt 2.2 ETL stycke 4 och 5 är väldigt lika. Stycke 4 ”Extraktionen kan göras med hjälp av JDBC (...) eller något annat databaskopplings-program...” och stycke 5 ”Extraktionen kan utföras på flera sätt.”. Finns det någon anledning till detta, eller är det att betraktas som återupprepning?

Figur 2: ETL-process skulle behöva en kortare förklaring för att underlätta förståelsen för denna figur då den är lite komplex.

Avsnitt 2.3 OLAP sidan 8 stycke 2 rad 7-9 ”...från datalager eller operativa system in i arbetsstationen...”, vad menas med operativa system?

Sidan 8 stycke 3 rad 13 ordet ”drill-through” saknar förklaring.

Sidan 8 stycke 4 rad 20 ordet ”aggregeringsnivåer” saknar förklaring.

Avsnitt 2.4 Integrering av OLAP-processen i ETL-systemet oklart vad författaren vill framhäva under avsnittet, behöver förtydligande.

2.3 Kapitel 3 Optimering

Avsnitt 3.3.1 Använda explain Plan och TKPROF när det är möjligt sidan 16 stycke 1 rad 10 ordet TKPROF saknar förklaring.

Avsnitt 3.3.3 Regelbaserad Optimering sidan 17 stycke 1 rad 1-3 ”Oracle slutar att ge stöd till denna modell i version 1 av Oracle10g, medan den stöds i version 2, för att det skall finnas bakåtkompatibilitet.” Behöver en omformulering då det är något rörigt just nu.

Avsnitt 3.3.5 Förstå sig på optimeraren detta avsnitt innehåller viktig information känns det som, dock så är det ganska svårt nu att utläsa denna information på ett bra sätt. Det skulle behövas klarare uppdelning mellan de olika sätten etc. Exempel på delar som är svårt att förstå är:

Rad 1 ”SQL kan exekveras med optimering baserad på regler eller kostnad.”

Rad 8-9 ”SQL kan endast optimeras med optimering baserad på kostnad”

Det känns som dessa två rader är motsägelsefulla.

Avsnitt 3.3.7 Urvalet är det kritiska i SQL-satsen sidan 18 stycke 1 rad 1-3 ”Generellt gäller att alla frågor som innehåller tecken som innehåller jämförelseoperatorer inte kommer

att indexeras” Denna mening behöver en omformulering, så att det tydligt framgår vad som menas.

Sidan 18 stycke 1 rad 3-4 ”Indexering kommer heller inte att ske vid kontroll om värde saknas(null) eller om värde inte saknas” Behöver ytterligare förklaring vad som menas med detta, är innebörden att indexering aldrig sker?

Sidan 19 de exempel som påvisar om indexering används eller inte skulle behöva någon liten ytterligare förklaring för att verkligen påvisa vad som menas/händer vid de båda olika fallen.

2.4 Kapitel 4 Befintligt system

Endast några småfel som återfinns med markering i uppsatsen.

2.5 Kapitel 5 Utveckling av mätare i QlikView

Endaste några småfel som återfinns med markering i uppsatsen.

2.6 Kapitel 6 Analys av batch/online

Avsnitt 6.2 Utvärdering av testkörning 1 sidan 27 stycke 1 rad 8 ordet paket behöver en närmare förklaring.

2.7 Kapitel 7 Slutsats och diskussion

Endaste några småfel som återfinns med markering i uppsatsen.

2.8 Övriga kommentarer

Småfel så som stavfel, grammatiska fel och så vidare återfinns direkt i rapporten och tas inte upp överhuvudtaget i denna oppositionsrapport.

3 Slutliga kommentarer

Uppsatsen är väl genomarbetad och bra. Uppsatsen har en klar ”röd tråd” som det inte är några problem att följa genom de olika kapitlen som läsare. Det känns som en bra blandning mellan bilder, informativ text och exempel.