

Medbit Data Warehouse and Data Mart

Document author: anna.hammais@tyks.fi

Preliminary plan: Steps of the ETL process

1. Aloitamme Opera-datasta (seuraavaksi varmaankin Toti ja Radu)
2. Haemme DataMartista Operaan liittyvät faktataulut (DMF_LEIKKAUS ja DMF_TOIMENPIDE) ja niihin liittyvät dimensiot
3. Haemme DW:stä taulun S_Toimenpiteet_Opera, josta saamme pre- ja postop-diagnoosit
4. Haku tehdään taulu kerrallaan meidän staging arealle. Teidän kannassa ei tarvitse tehdä joineja tms. eikä teidän serveri kuormitu varmaan kummemmin.
5. DW:ssä käytetään act_id-sarakkeita pääavaimina ja latpvm-saraketta tiedon uutuuden määrittämiseen.
6. DM:ssä käytetään pääavaimena taulun PK-saraketta ja tiedon uutuuden määrittämiseen sarakkeita DATE_INSERT ja DATE_UPDATE.
7. DW-tasolla pitää aina huomioida valid kentän arvo. Y = viimeisin voimassaoleva (rivi voi silti olla lähteessä poistettu tai peruttu), N = tietovarastosta löytyy riville uudempi rivi, jossa joku arvoista on päivittynyt. Hakea vain valideja tietoja joka kohdassa, eli tiedon täytyy olla validi DW:ssä ja lähteessä, ei riitä että jommassakummassa.
8. Kun alkulataus on tehty, uudet ja muuttuneet tiedot haetaan esim. kerran yössä. Tarkistetaan ennen jokaisen taulun hakua, että teidän latauskierros on valmis, taulusta VarSat_SA_VSSHP.SA_VSSHP_logTable (etlJobName = 'JOB_Tietovarasto_VSSHP' and status = 'ONNISTUI' and endTime is not null).

Kohta 7 ontuu vielä, koska systeemitunnuksellamme SVCL_KTPLukija ei ole lukuoikeutta VarSat_SA_VSSHP-kantaan (vain HAMMAISA on oikeus).