Udskrevet er dokumentet ikke dokumentstyret. **The ETL-schema**					DSB
Dokumentbrugere: **KONC-POEM-MT**	Dokumentansvarlig: **D-PØMQansvar**	Redaktør: Version: **1.1**	Dokumentnummer: **12. 6. 1.05**	Godkendt af: **PAMA2610** **08.03.2010**	Niveau:

1) Introduction

The ETL process consists of extract data (capture changes), check, clean, transform, conform and deliver data to the DM structures and ABM model, such as, dimensions, fact tables, hierarchies, etc.

This schema keeps metadata of the ETL processes as well as stored procedures called by the ETL packages or processes. Therefore, the tables in this schema are used to log information about the ETL process.

2) Prefixes

To identify whether the table keeps data related to the ODS or EDW schema or general data not related to a particular schema, the following prefixes are used:

Prefix	Description
Grl_	Log data which is not related to a particular schema
Ods_	Log data related to the ODS schema
Edw_	Log data related to the EDW schema

5) Tables

The tables are accesed through different stored procedures on the MDW-database. Refer to these for details about parameters and datatypes.

5.01) etl.grl_PackageLog

This table registers data about packages executed in the system, such as, the execution time, a flag indicating whether the package was executed successfully, etc.

Etl.grl_PackageLog table description:

Column name	Description
PackageLogID	ID assigned to the package by the system
PackageName	Packagename
StartTime	Package starttime
EndTime	Package endtime
Success	Flag indicating whether the package was executed successfully or not, using two values, True=Succesful, False=Unsuccessful

5.02) etl.ods_LoaderLog

This table stores information about the loading of data from the source systems to the ODS schema. The table contains the following columns:

Etl.ods_LoaderLog table description:

Column name	Description	Example	Comment
LoaderLogID	Loader unique id	1,2	Sequential number assigned by the system. There could be more than one loader per package
PackageLogID	Package unique id	1,2	Packageid passed as a parameter
SourceSystem	Short name of the source system whose data is loaded	PDS, APL	Refer to the list in section 12.06.01.1, for a full list of source system short names
TableName	Name of the table being loaded in the package.	ods.RD_LTD2_Lokopersonale, ods.RD_APL_Forbrug	**Note: Table name includes scheme prefix**
StartTime	Date and time when the data loader starts	10-09-2008 14:14:40	Field format includes date and time
EndTime	Date and time when the data loader stops	10-09-2008 14:16:00	Field format includes date and time
Success	Flag indicating whether the data load was executed successfully or not, using two values, True = Successful, False = Unsuccessful	True, False	
MaxPeriod	Holds the latest period of the loaded data.	200803, 2007Q2	Example: If data is from jan, feb, mar 2008 the latest period would be mar 2008.
NumRows	Column containing the Number of Rows loaded from the data source	200	Number indicating the number of rows loaded
CtlColName_1	Name of the control column 1, whose amounts are aggregated and shown in the column CtlColTot_1	Antal minutter	
CtlColTot_1	Total sum of column whose name is shown in CtlColName_1	1234567	
The following columns are used only when there is more than one control column in the file			
CtlColName_2	Name of the control column 2, whose amounts are aggregated and shown in the column CtlColTot_2	Antal minutter	
CtlColTot_2	Total sum of column whose name is shown in CtlColName_2	1234567	
CtlColName_3	Name of the control column 3, whose amounts are aggregated and shown in the column CtlColTot_3	Antal minutter	
CtlColTot_3	Total sum of column whose	1234567	

	name is shown in CtlColName_3		
CtlColName_4	Name of the control column 4, whose amounts are aggregated and shown in the column CtlColTot_4	Antal minutter	
CtlColTot_4	Total sum of column whose name is shown in CtlColName_4	1234567	

5.03) etl.edw_TabRefLog

This table keeps data related to EDW datatables refresh processes.

Etl.edw_TabRefLog table description:

Column name	Description	Example	Comment
TabRefLogID	Table refresh task unique identifier	1,2	Sequential number assigned by the system. There could be more than one table refresh tasks per package
PackageLogID	Package unique id	1,2	Packageid passed as a parameter
TableName	Name of the updated EDW table	GD_LTD_Togpersonale	
TypeRefresh	Code indicating how data is refreshed	T/I= Truncate and Insert, INS = Insert, DEL = Delete, UPD = Column update	This code refers to the refresh strategy. For a full list of these codes, refer to the EDW schema
Period	Period from which the data is being refreshed	200801	This period is taken by the data periods contained in the RD table
NumRowsDel	If data is discarded from the table, this cells can be used to keep the number of discarded rows	200	Number of rows to be discarded
NameCtlColDel	If data is discarded, this cell storages the name of the main control column whose totals will be aggregated and kept as a reference	Turtimer	
TotCtlColDel	If data is discarded, this cell storages the aggregated amount of the discarded data, whose name is shown in ColNameDel	100000	
NumRows	Number of rows inserted in the table. If RefreshType = UPD, this cell shows the total rows being updated. If RefreshType = DEL, this column is NULL	200	
RefColName_1	Name of the control column 1, whose amounts are aggregated and shown in the column RefColTot_1	Turtimer	
RefColTot_1	Total sum of column whose name is shown in RefColName_1	2000000	
RefColName_2 **	Name of the control column 2, whose amounts are aggregated and shown in the column RefColTot_2		
RefColTot_2 **	Total sum of the column whose name is shown in RefColName_2		

RefColName_3 **	Name of the control column 3, whose amounts are aggregated and shown in the column RefColTot_3		
RefColTot_3 **	Total sum of the column whose name is shown in RefColName_3		
RefColName_4 **	Name of the control column 4, whose amounts are aggregated and shown in the column RefColTot_4		
RefColTot_4 **	Total sum of the column whose name is shown in RefColName_4		
RefColName_5 **	Name of the control column 5, whose amounts are aggregated and shown in the column RefColTot_5		
RefColTot_5 **	Total sum of the column whose name is shown in RefColName_5		
StartTime	Date and time when the refresh task starts	10-09-2008 14:14:40	Format includes date and time
EndTime	Date and time when the refresh task stops	10-09-2008 14:16:00	Format includes date and time
Sucess	Flag indicating whether the data refresh was executed successfully or not, using two values, True = Successful, False = Unsuccessful	True, False	

** These columns are used only when there is more than one control column in the file

Revisioner

Version	Godkendt	Revisions information
1	13.02.2009	
1.1	08.03.2010	