

Datawarehouse Et Data Lake

Documentation technique

28/05/2024



→ Présentation

- Datawarehouse p.4
- Data Lake p.5
- Objectifs p.6

→ Liste et description des tables du Datawarehouse p.7

→ Liens entre les tables de fait et les tables de dimension p.10

→ Liste et description des tables du Data Lake p.23

→ Liens entre les tables du Datawarehouse et Data Lake p.27

- **Cette documentation a pour objectif de vous fournir une vue d'ensemble du Datawarehouse et du Data Lake que propose Peopulse.**
- **Elle a également pour objectif de vous fournir une compréhension de la structure du Datawarehouse et du Data Lake, en présentant la liste et la description des tables qui la composent, ainsi que les relations entre ces tables en mettant en évidence les clés primaires qui les lient.**

Datawarehouse Peopulse



Le « Datawarehouse » Peopulse est un environnement de base de données dédié à un client. Il recueille des données de sources variées contenant les informations liées aux contrats, aux demandes, aux relevés d'heures, aux préfatures et aux factures. **Le Datawarehouse contient des données agrégées, pas les données brutes.**

La structure en étoile est l'architecture utilisée dans le datawarehouse. Au cœur de la structure en étoile se trouve des tables de faits et des tables de dimensions.

Les tables de faits vont contenir des mesures numériques ou quantitative du recours à l'intérim telles que le nombre de demandes, le nombre de contrats, le volume d'heures ou encore le montant des factures.

Les tables de dimensions vont contenir des informations contextuelles et des attributs des données qui fournissent des détails sur les mesures stockées dans les tables de faits. Elles contiennent des informations descriptives comme les clients, les agences, les différents statuts... Elles facilitent la compréhension textuelle des données des tables de faits.

Le datawarehouse ne contient de données personnelles.

Data Lake Peopulse



Le « Data Lake » Peopulse est un emplacement de stockage centralisé qui contient toutes les datas d'un client sous un format brut et granulaire. **Il contient les données brutes qui sont agrégées dans le datawarehouse.**

Le Data Lake contient des données personnelles.

→ **Le Data Warehouse et le Data Lake ont été conçu pour répondre aux besoins suivants :**

- Centralisation des données
- Consolidation des données
- Amélioration de la performance des requêtes
- Intégration avec des outils d'analyse pour les prises de décisions stratégiques

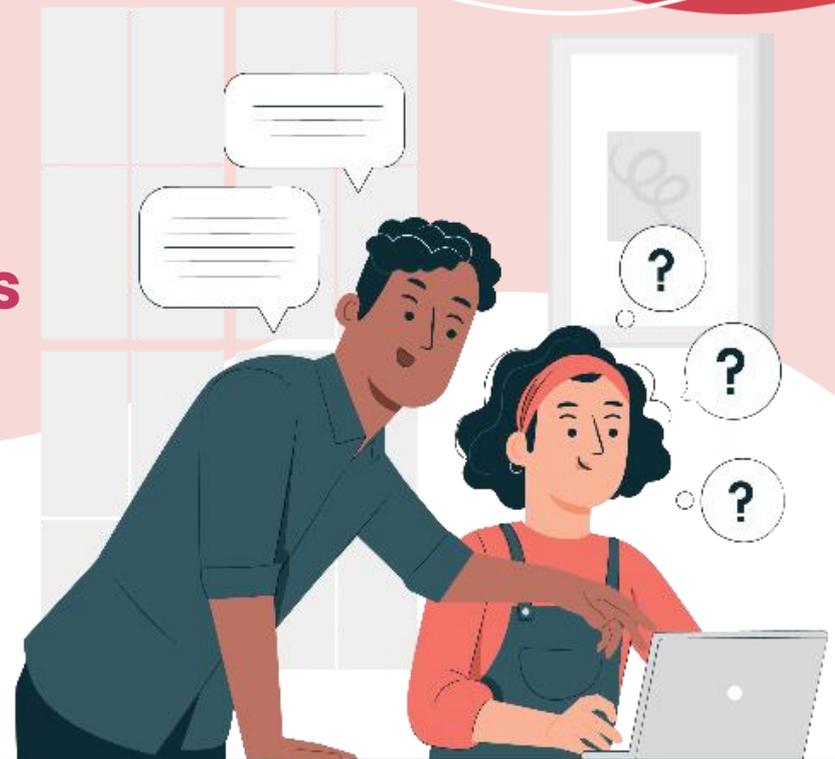
→ **Le Data Warehouse et le Data Lake sont complémentaires et peuvent être utilisés ensemble pour répondre aux besoins variés d'analyse des données.**

→ **Les données structurées et agrégées provenant du Datawarehouse peuvent être mise en relation avec les données du Data Lake afin d'enrichir l'analyse avec des données brutes.**

→ **Ensemble, le Datawarehouse et le Data Lake offrent une solution complète pour la gestion et l'analyse de données.**



Liste et description des tables du Datawarehouse



Les tables de faits

Nom des tables	Contenu des tables
demande_initiale	regroupe toutes les informations concernant les demandes (date de création, quantité demandée, date de la 1 ^{ère} réponse, date de début, date de fin...)
contrat	regroupe toutes informations concernant les contrats (date d'intégration, date de signature EU, date de refus...)
mission	regroupe les informations liées aux missions (date de début, date de fin, date de souplesse min, date de souplesse max...)
pointage	regroupe les informations liées aux heures (période, quantité d'heure, montant, site client, agence ETT...)
préfacture	regroupe les informations liées aux préfactures (numéro de préfacture, montant, site client, agence ETT...)
facture	regroupe toutes les informations concernant les factures (numéro de facture, montant, numéro de préfacture, écart avec la préfacture, site client, agence ETT...)

Les tables de dimensions

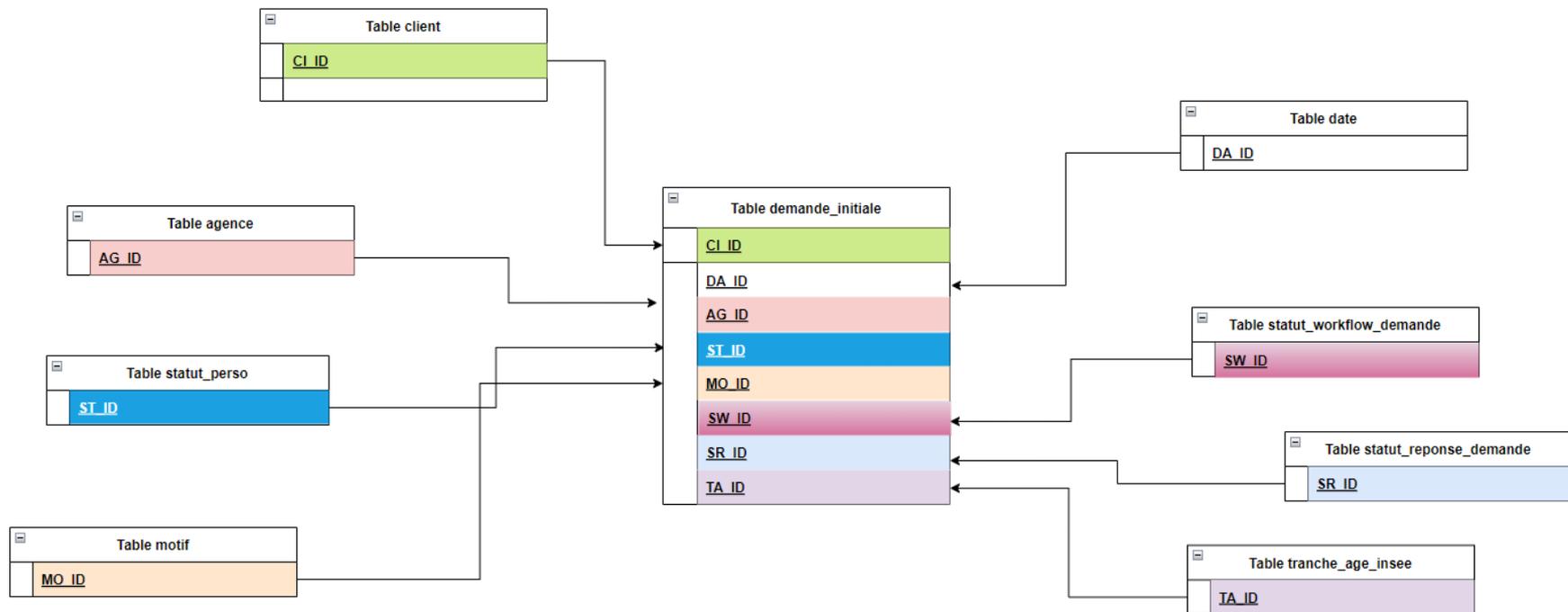
Nom des tables	Contenu des tables
agence	regroupe toutes les informations concernant les agences (nom de l'agence ETT, nom de l'enseigne ETT...)
client	regroupe toutes informations concernant les sites clients (nom du site, ID du site, code analytique et sous-analytique, code ERP...)
date	regroupe les informations liées aux dates (jour, mois, année, semaine...)
motif	regroupe les informations liées aux motifs de recours (ID du motif, code du motif, libellé du motif)
statut_perso	regroupe les informations liées aux statuts des intérimaires (libellé sexe, code sexe, statut handicap, ID du contrat...)
statut_workflow_demande	regroupe les informations concernant le statut du workflow dans les demandes (ID statut, libellé du workflow)
statut_reponse_demande	regroupe les informations concernant liées au statut de réponse d'une demande (ID statut, libellé du statut)
tranche_age_insee	regroupe les informations concernant liées aux tranches d'âge (ID de la tranche d'âge, libellé de la tranche d'âge)



Lien entre les tables de faits et les tables de dimensions



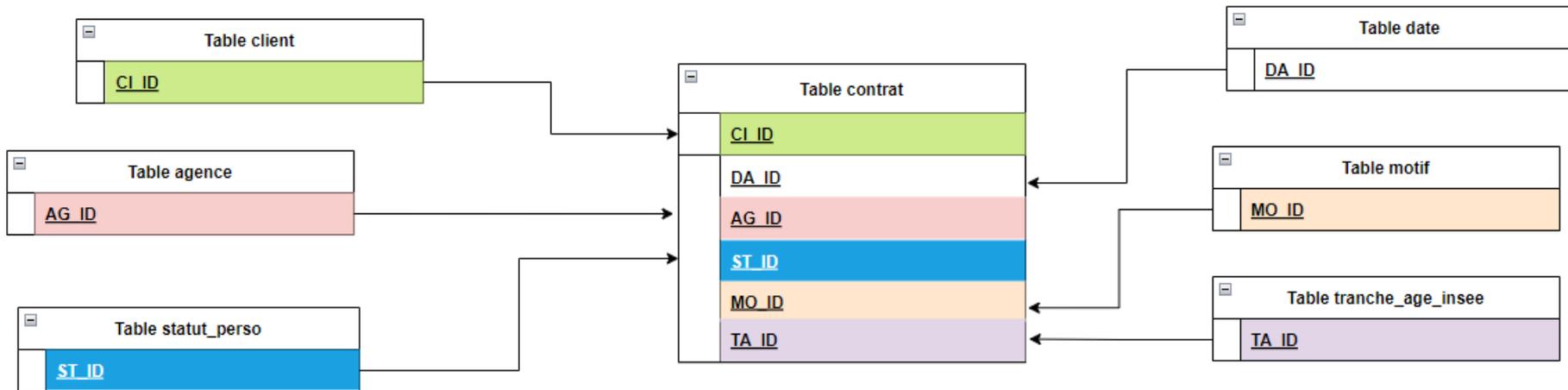
Jointure avec la table demande_initiale



Jointure avec la table demande_initiale

Tables	Tables	Clés
client	demande_initiale	cl_id
date	demande_initiale	da_id
agence	demande_initiale	ag_id
statut_perso	demande_initiale	st_id
workflow_demande	demande_initiale	sw_id
statut_reponse_demande	demande_initiale	sr_id
tranche_age_insee	demande_initiale	ta_id
motif	demande_initiale	mo_id

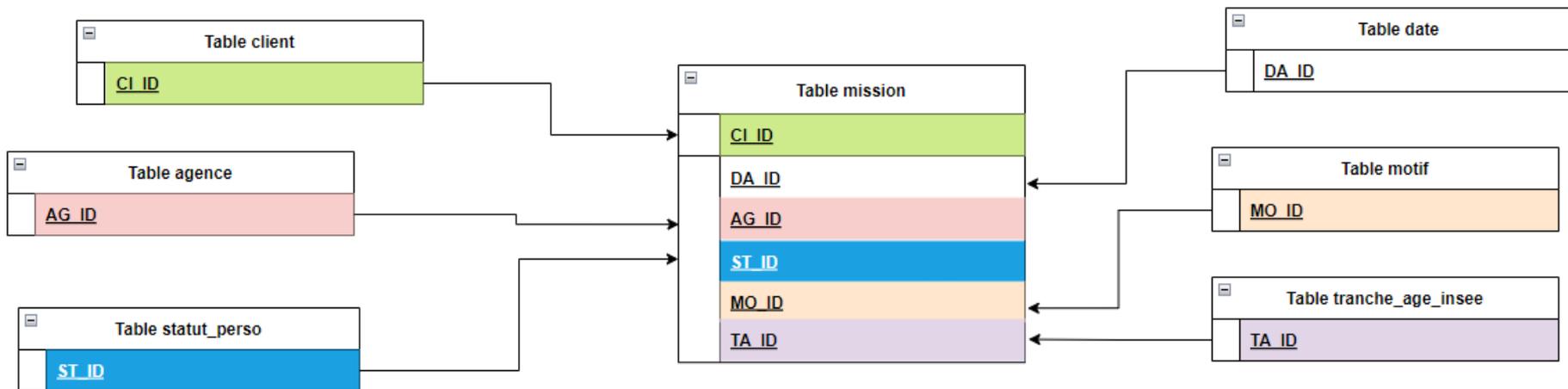
Jointure avec la table contrat



Jointure avec la table contrat

Tables	Tables	Clés
client	contrat	cl_id
date	contrat	da_id
agence	contrat	ag_id
statut_perso	contrat	st_id
motif	contrat	mo_id
tranche_age_insee	contrat	ta_id

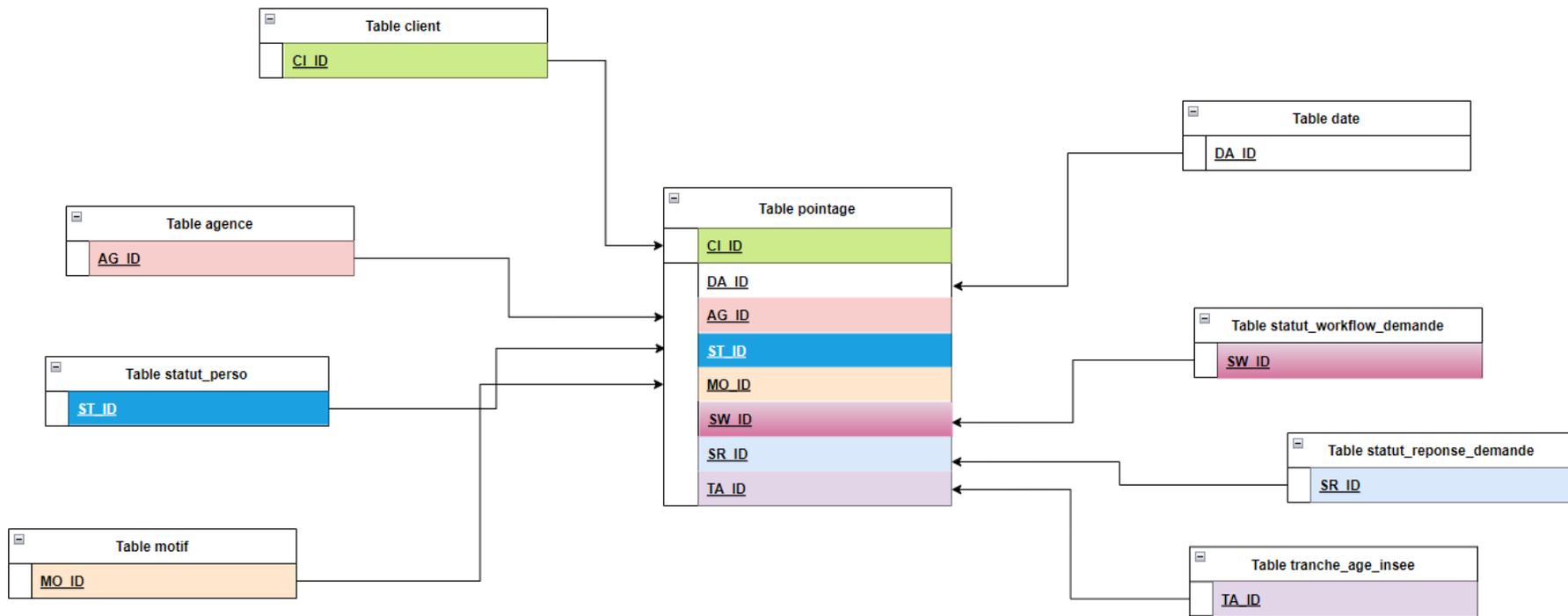
Jointure avec la table mission



Jointure avec la table mission

Tables	Tables	Clés
d_client	mission	cl_id
d_date	mission	da_id
d_agence	mission	ag_id
d_motif	mission	mo_id
d_statut_perso	mission	st_id
d_tranche_age_insee	mission	ta_id

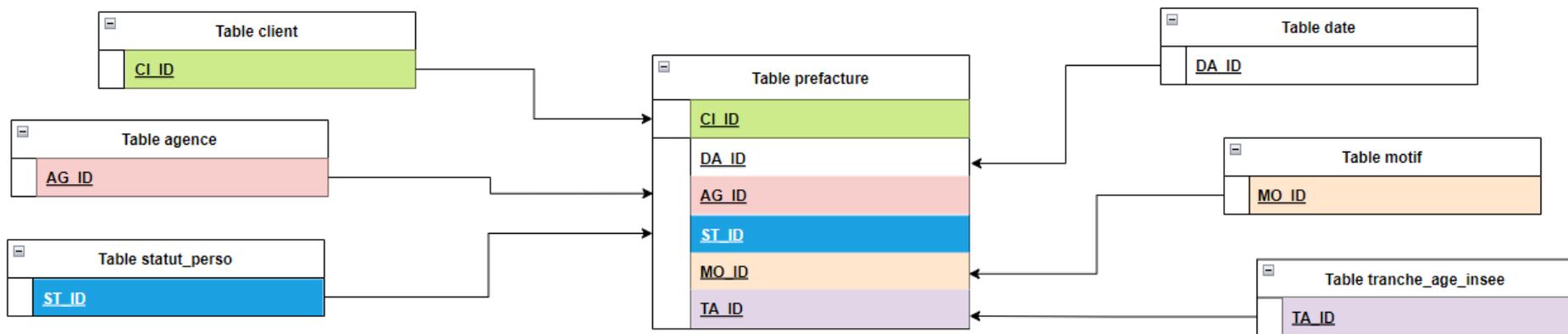
Jointure avec la table pointage



Jointure avec la table pointage

Tables	Tables	Clés
d_client	pointage	cl_id
d_date	pointage	da_id
d_agence	pointage	ag_id
d_statut_perso	pointage	st_id
d_workflow_demande	pointage	sw_id
d_statut_reponse_demande	pointage	sr_id
d_tranche_age_insee	pointage	ta_id
d_motif	pointage	mo_id

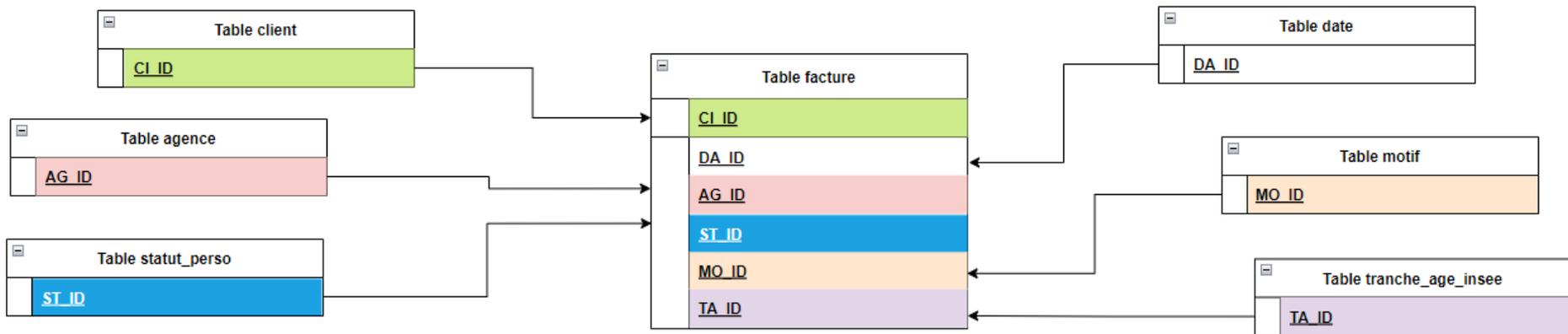
Jointure avec la table prefature



Jointure avec la table facture

Tables	Tables	Clés
d_client	prefacture	cl_id
d_date	prefacture	da_id
d_agence	prefacture	ag_id
d_tranche_age_insee	prefacture	ta_id
d_motif	prefacture	mo_id
d_statut_perso	prefacture	st_id

Jointure avec la table facture

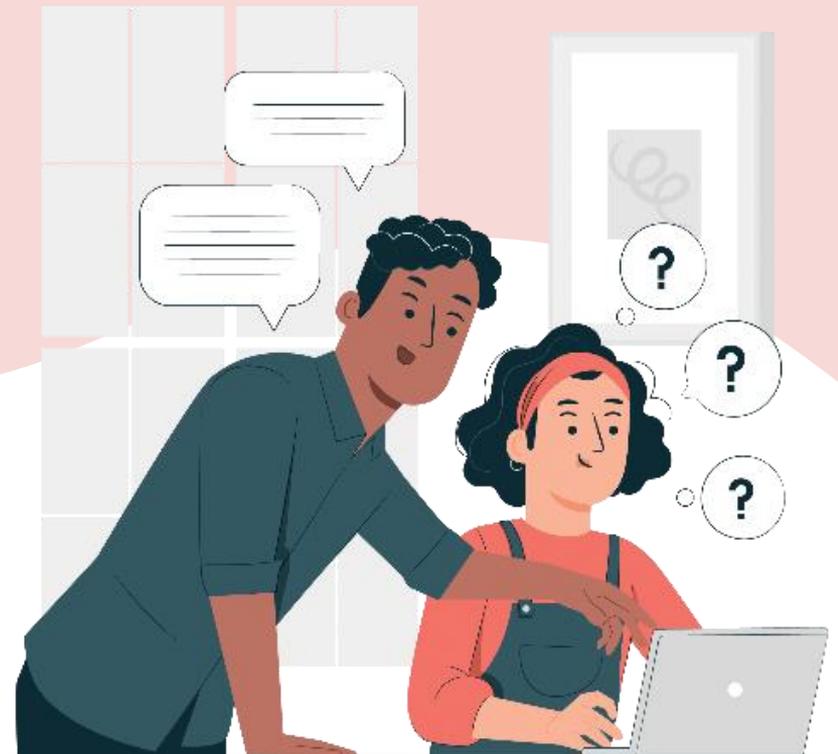


Jointure avec la table facture

Tables	Tables	Clés
client	facture	cl_id
date	facture	da_id
agence	facture	ag_id
tranche_age_insee	facture	ta_id
motif	facture	mo_id
statut_perso	facture	st_id



Liste et description des tables du Data Lake



Liste des tables

Nom des tables	Contenu des tables
dl_agence	regroupe toutes les informations concernant les agences ETT (adresse, Siret, nom, téléphone, adresse e-mail...)
dl_client	regroupe toutes informations concernant les sites d'un client (Identifiant du site, adresse postale, ville, Siret, adresse e-mail, téléphone, actif/non actif, mode de rupture...)
dl_mission	regroupe les informations liées aux contrats de mise à disposition (n° du contrat, date de début et date de fin de mission, qualification, lieu de travail, site client, ETT...)
dl_mission_avenant	regroupe les informations liées aux avenants de contrat (n° d'avenant, date de début et fin...)
dl_mission_rubrique	regroupe les informations liées aux rubriques des missions (libellé, code, taux payé, taux facturé, coefficient...)
dl_client équipe	regroupe toutes les informations concernant les équipes d'un client (libellé de l'équipe, code équipe, rubriques liées à l'équipe, base horaire, horaire...)
dl_demande	regroupe les informations liées aux demandes (n° de la demande, statut de la réponse, date de création de la demande, date d'envoi de la demande, durée de la mission, motif, justification...)

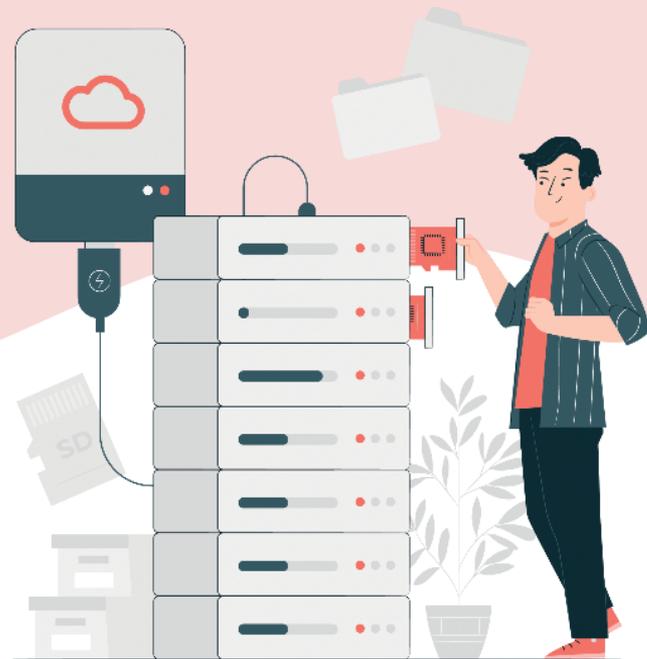
Liste des tables

Nom des tables	Contenu des tables
dl_demande_agence	regroupe les informations liées aux candidatures proposées par les ETT (nombre d'intérimaire proposé, nb d'intérimaire accepté, ranking de l'ETT...)
dl_demande_candidat	regroupe les informations liées aux statuts des candidats dans les demandes (nom, prénom...)
dl_personnel	Regroupe les informations liées aux intérimaires (nom, prénom, matricule ETT, matricule EU, coordonnées, date de naissance, habilitation...)
dl_contrat	regroupe toutes les informations liées aux contrats (numéro de contrat, date de début...)
dl_penibilité	regroupe les informations liées aux pénibilités des demandes et contrats (libellé de la pénibilité, code de la pénibilité...)
dl_prerequis	regroupe les informations liées aux prérequis des demandes et contrats (libellé du prérequis, code du prérequis...)
dl_qualification	regroupe les informations liées aux qualifications (libellé de la qualification, code de la qualification...)

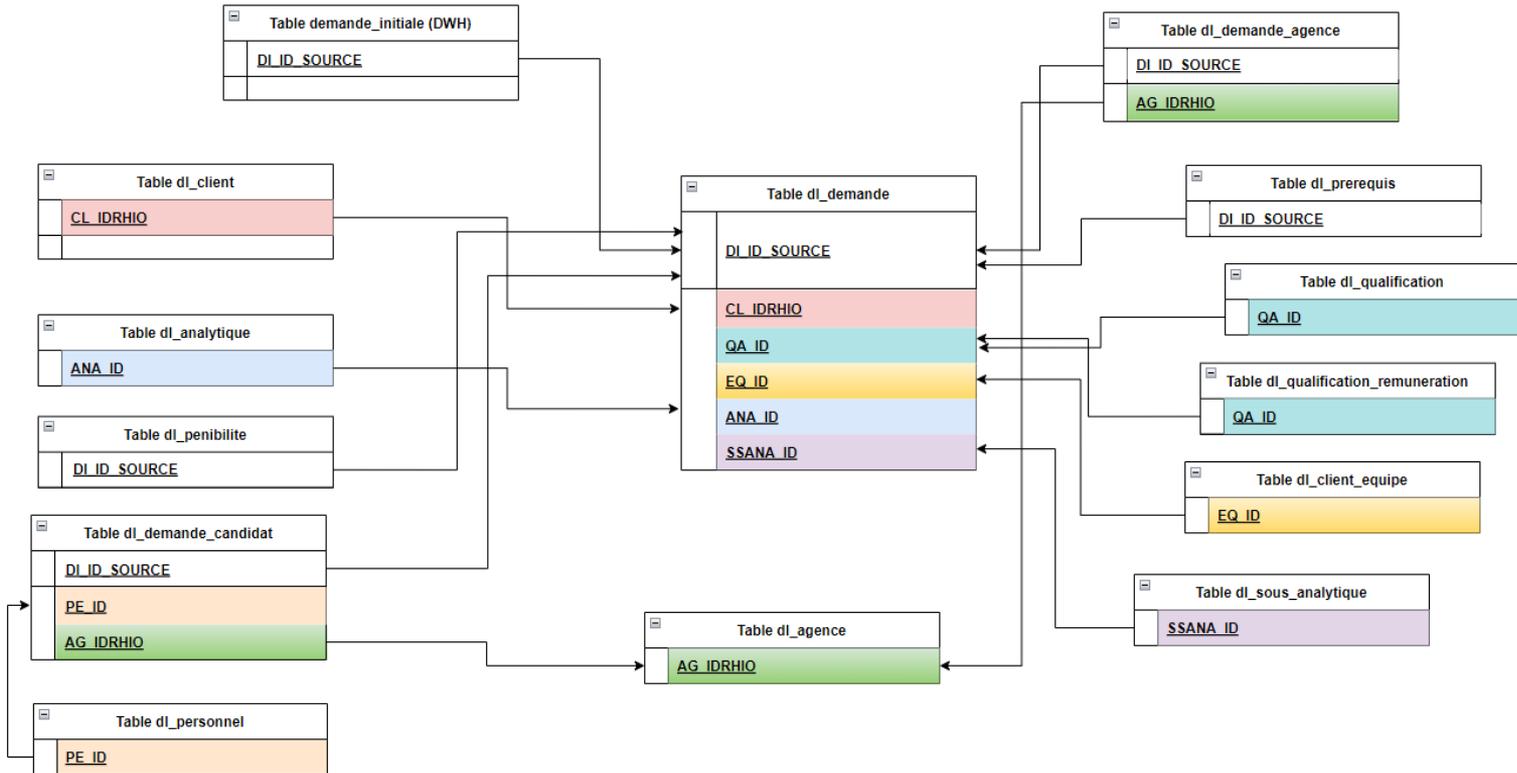
Tables	Tables
dl_qualification_rémunération	regroupe les informations liées aux rémunérations d'une qualification
dl_rubrique	regroupe les informations liées aux rubriques paramétrées dans la plateforme (code, libellé, taux, valorisation...)
dl_analytique	regroupe les informations liées aux analytiques (code analytique, libellé analytique, code ERP...)
dl_sous_analytique	regroupe les informations liées aux sous-analytiques (code sous-analytique, libellé sous-analytique, code ERP...)
dl_pointage	regroupe les informations liées aux détails des heures des intérimaires (n° de semaine, qualification, site, nb d'heures effectuées, quantité des rubriques...)
dl_préfecture	regroupe les informations liées aux préfectures (n° de préfecture, date, statut de la préfecture, montant...)
dl_facture	Regroupe les informations liées aux factures (n° de facture, type, date, site, ETT, statut de la facture, montant HT, montant TVA, montant TTC...)



Lien entre les tables du Datawarehouse et du Data Lake



Module "demandes"



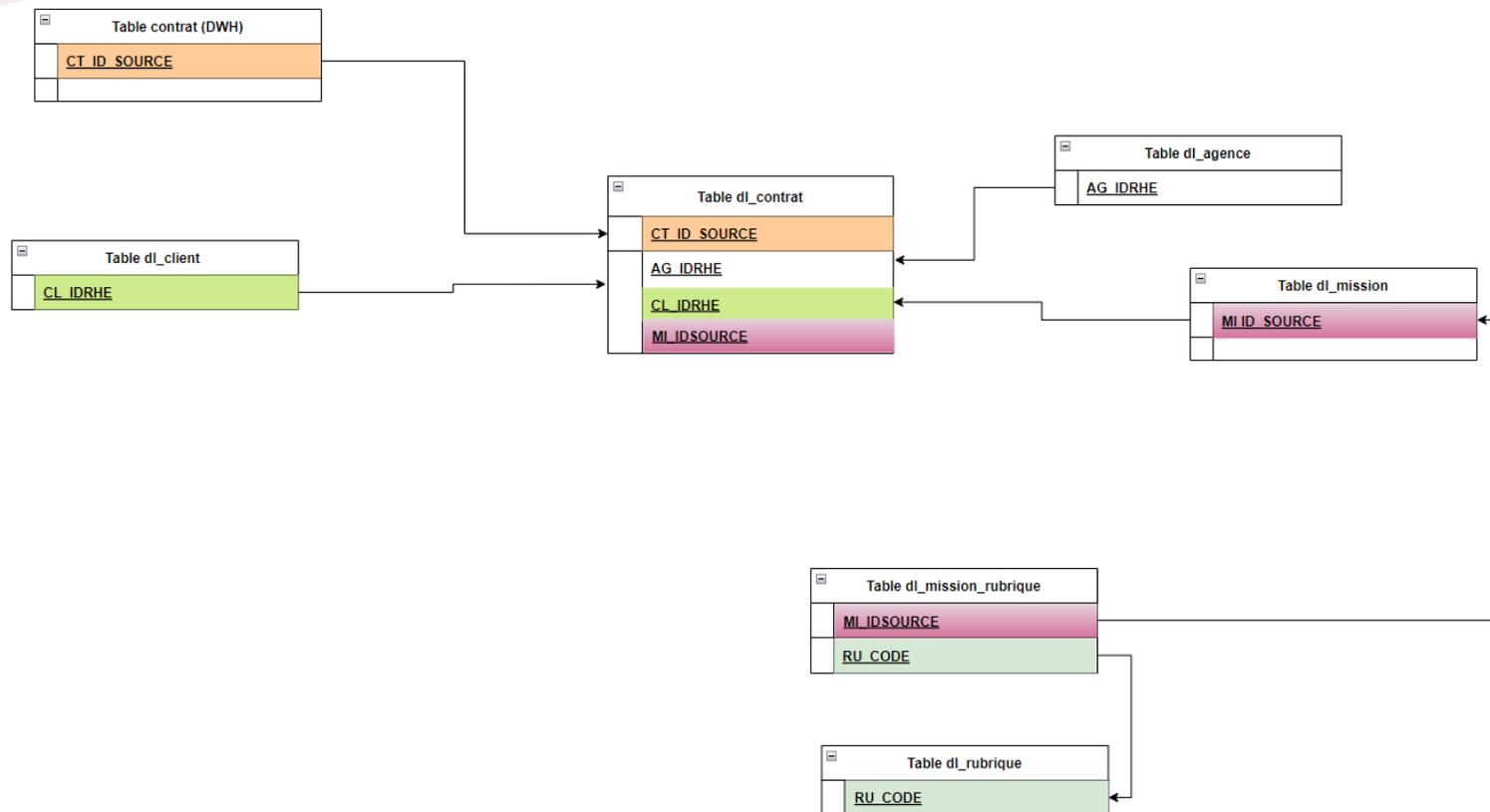
Jointure entre les tables

Tables	Tables	Clés
demande_initiale (DWH)	dl_demande	di_id_source
dl_client	dl_demande	cl_idrhio
dl_penibilite	dl_demande	di_id_source
dl_demande_agence	dl_demande	di_id_source
dl_demande_agence	dl_agence	ag_idrhio
dl_prerequis	dl_demande	di_id_source
dl_qualification	dl_demande	qa_id

Jointure entre les tables

Tables	Tables	Clés
dl_qualification_remuneration	dl_demande	qa_id
dl_client_equipe	dl_demande	eq_id
dl_demande_candidat	dl_demande	di_id_source
dl_demande_candidat	dl_personnel	pe_id
dl_demande_candidat	dl_agence	ag_idrhio
dl_analytique	dl_demande	ana_id
dl_sous_analytique	dl_demande	ssana_id

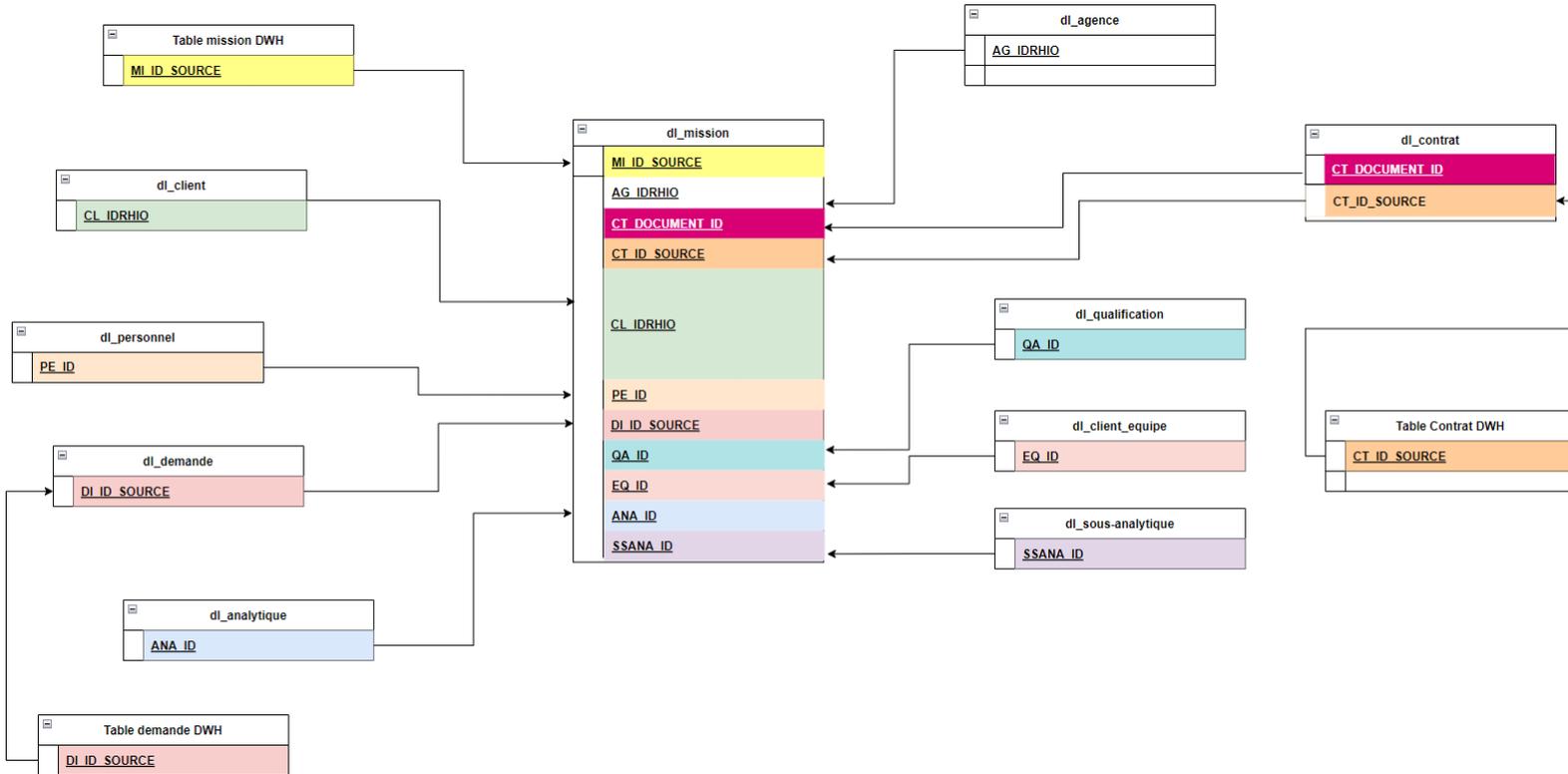
Module "contrats"



Jointure entre les tables

Tables	Tables	Clés
contrat (DWH)	dl_contrat	ct_id_source
dl_agence	dl_contrat	ag_idrhe
dl_client	dl_contrat	cl_idrhe
dl_mission	dl_contrat	mi_id_source
dl_mission_rubrique	dl_mission	mi_id_source
dl_mission_rubrique	dl_rubrique	ru_code

Module "missions"



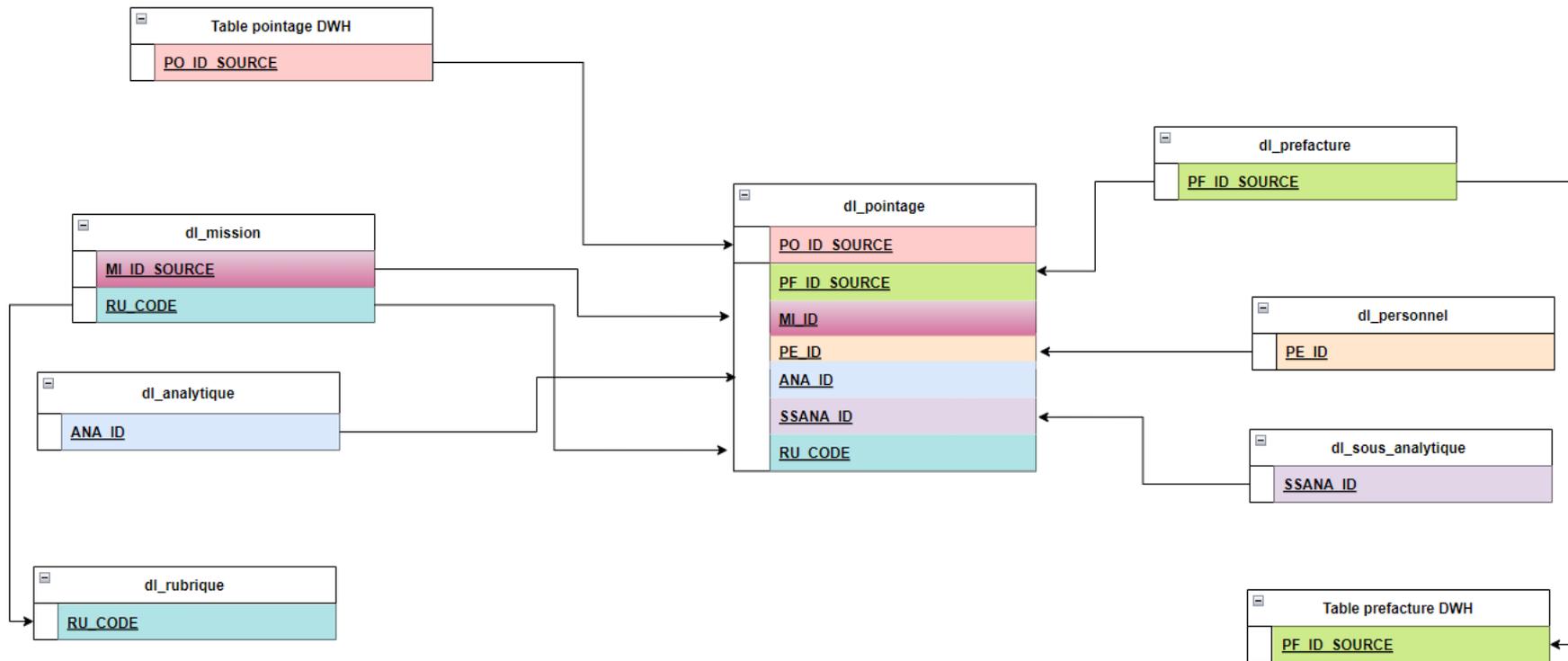
Jointure entre les tables

Tables	Tables	Clés
mission (DWH)	dl_mission	mi_id_source
dl_agence	dl_mission	ag_idrhio
dl_client	dl_mission	cl_idrhio
dl_contrat	dl_mission	ct_document_id
dl_contrat	dl_mission	ct_id_source
dl_personnel	dl_mission	pe_id
dl_demande	dl_mission	di_id_source
dl_demande	demande_initiale (DWH)	di_id_source

Jointure entre les tables

Tables	Tables	Clés
dl_analytique	dl_mission	ana_id
dl_qualification	dl_mission	qa_id
dl_client_equipe	dl_mission	eq_id
dl_sous_analytique	dl_mission	ssana_id
dl_contrat	contrat (DWH)	ct_id_source

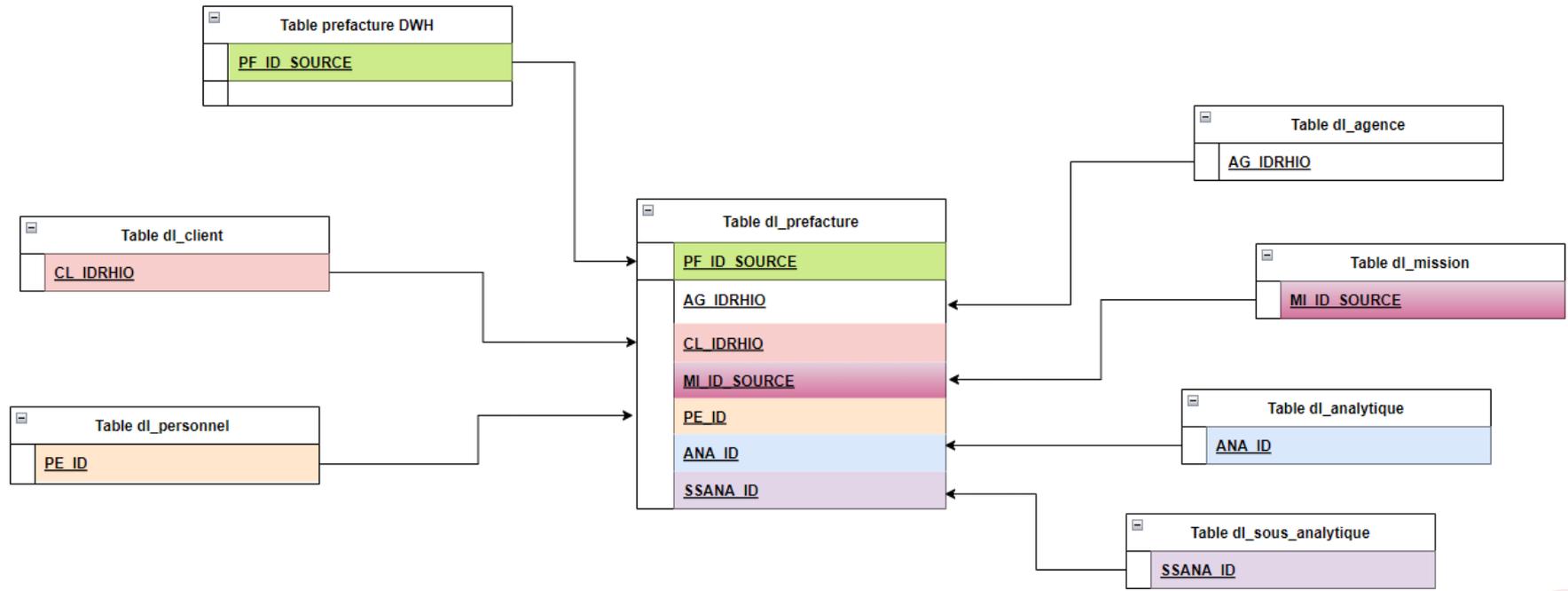
Module "heures"



Jointure entre les tables

Tables	Tables	Clés
pointage (DWH)	dl_pointage	po_id_source
dl_prefacture	dl_pointage	pf_id_source
prefacture (DWH)	dl_pointage	pf_id_source
dl_mission	dl_pointage	mi_id_source
dl_mission	dl_pointage	ru_code
dl_mission	dl_rubrique	ru_code
dl_pointage	dl_personnel	pe_id
dl_analytique	dl_pointage	ana_id
dl_sous_analytique	dl_pointage	ssana_id

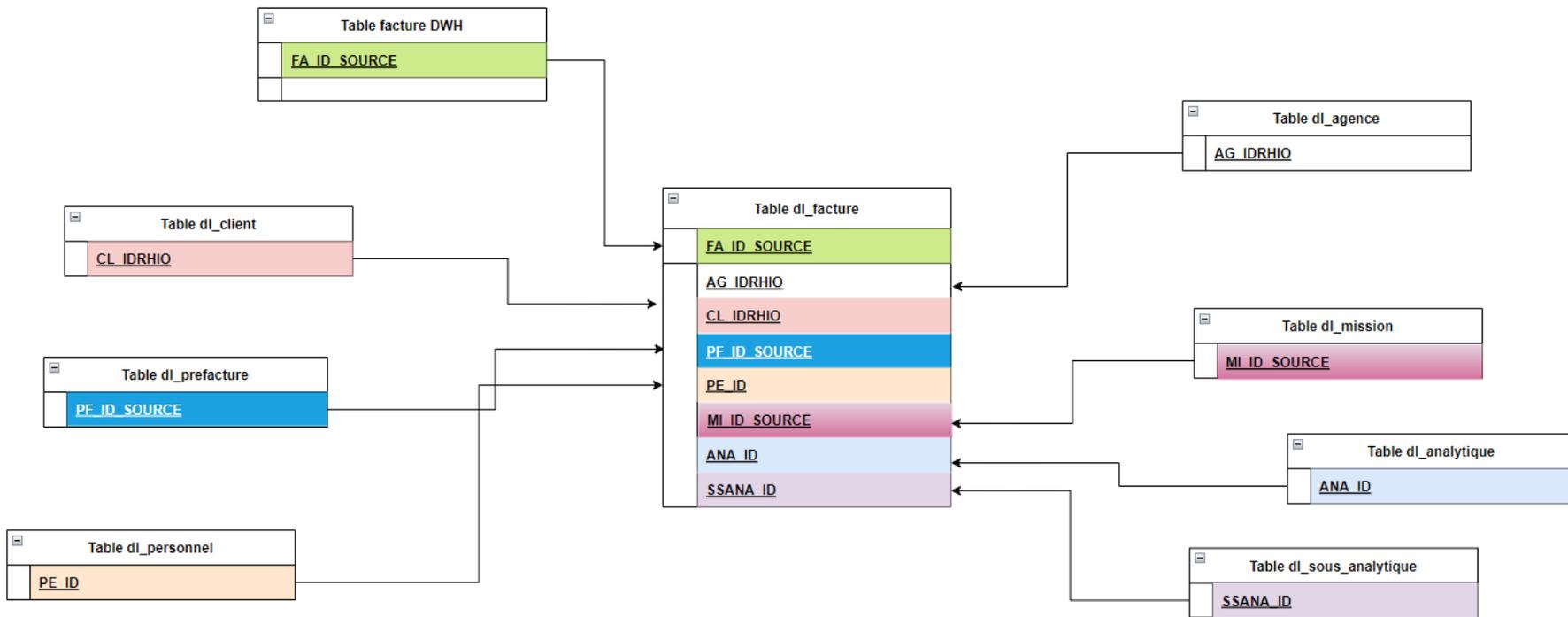
Module "prefactures"



Jointure entre les tables

Tables	Tables	Clés
prefecture (DWH)	dl_prefecture	pf_id_source
dl_client	dl_prefecture	cl_idrhio
dl_personnel	dl_prefecture	pe_id
dl_agence	dl_prefecture	ag_idrhio
dl_mission	dl_prefecture	mi_id_source
dl_analytique	dl_prefecture	ana_id
dl_sous_analytique	dl_prefecture	ssana_id

Module "factures"



Jointure entre les tables

Tables	Tables	Clés
facture (DWH)	dl_facture	fa_id_source
dl_agence	dl_facture	ag_idrhio
dl_client	dl_facture	cl_idrhio
dl_prefacture	dl_facture	pf_id_source
dl_personnel	dl_facture	pe_id
dl_analytique	dl_facture	ana_id
dl_sous_analytique	dl_facture	ssana_id

Jointure entre les tables

Tables	Tables	Clés
facture (DWH)	dl_facture	fa_id_source
dl_analytique	dl_facture	ana_id
dl_analytique	dl_facture	ana_code
dl_sous_analytique	dl_facture	ssana_id
dl_sous_analytique	dl_facture	ssana_code

MERCI