

Rétrospective du projet Business Intelligence

Projet collaboratif à 3 : Méthodologie complète de bout en bout

■ **Objectif global** : Produire un Datawarehouse robuste & un ensemble de dashboards répondant à nos problématiques métiers.

■ **Durée** : 2 itérations sur 3 semaines (14 jours effectifs)

■ **Stack technique** : PostgreSQL, DBT, Power BI, Python

■ **Livrables** : DWH structuré, documentation, dashboard Power BI, présentation orale

Approche générale du projet



**NOTRE DÉMARCHE
COLLABORATIVE :**



**■ ANALYSE
COLLECTIVE DES
JEUX DE DONNÉES
(CSV)**



**■ TRAVAUX EN
BINÔME PUIS EN
TRINÔME SELON LES
ÉTAPES**



**■ CO-CONCEPTION
DU MODÈLE MÉTIER
ET DES BESOINS
ANALYTIQUES**



**■ RÉPARTITION DES
TÂCHES : INGESTION,
MODÉLISATION DBT,
VISUALISATION
POWER BI**



**■ SYNCHRONISATION
QUOTIDIENNE &
VALIDATIONS
INTERMÉDIAIRES**

Storytelling

CONTEXTE

■ **Claire Dubois**, chargée de mission territoriale, vient de recevoir une mission délicate : dresser un état des lieux du tissu économique de l'Isère. Son objectif est de comprendre comment les entreprises du département s'organisent, selon leur taille et leur secteur, et quelles dynamiques en ressortent.



Compréhension des données & problématique



Compréhension
des données &
problématiques



Étapes réalisées :



 Exploration
des fichiers CSV




 Vérification
des formats,
encodages, types



 Détection des
limites : valeurs
manquantes,
doublons



 Formulation
des
problématiques
métiers
(exemples) :

Comment se répartissent
les entreprises selon les
secteurs ?

Quels secteurs sont les
plus représentés par
commune ?

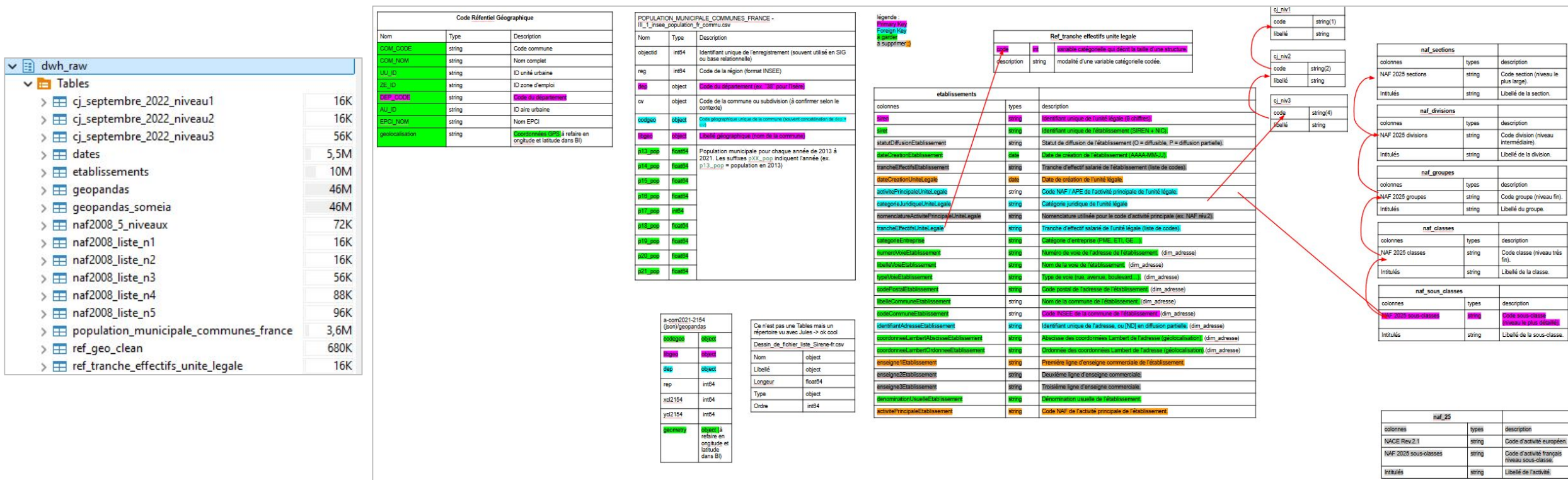
Comment évoluent les
catégories juridiques ?

Quelle concentration
d'entreprises par
territoire ?

*"Comment se structure
le tissu économique de l'Isère
?"*

Construction du DWH en 3 étapes

1. RAW



Construction du DWH en 3 étapes

2. STAGING

dwh_staging		
Tables		
> stg_cadre_juridique_niveau1		16K
> stg_cadre_juridique_niveau2		16K
> stg_cadre_juridique_niveau3		40K
> stg_dates		3,6M
> stg_etablissements		5M
> stg_geo		46M
> stg_naf_5_niv		72K
> stg_naf_n1		16K
> stg_naf_n2		16K
> stg_naf_n3		40K
> stg_naf_n4		72K
> stg_naf_n5		72K
> stg_population		3M
> stg_ref_geo		136K
> stg_tranches_effectifs		16K

Dim_commune

POPULATION MUNICIPALE COMMUNES FRANCE - Ill_1_insee_population_fr_commu.csv		
Nom	Type	Description
objectid	int64	Identifiant unique de l'enregistrement (souvent utilisé en SIG ou base relationnelle)
codegeo	objet	Code géographique unique de la commune (souvent concaténée de l'INSEE et du département)
libelle	string	Libellé géographique (nom de la commune)
p13_pop	float64	Population municipale pour chaque année de 2013 à 2021. Les suffixes pXX_pop indiquent l'année (ex. p13_pop = population en 2013)
p14_pop	float64	
p15_pop	float64	
p16_pop	float64	
p17_pop	float64	
p18_pop	float64	
p19_pop	float64	
p20_pop	float64	
p21_pop	float64	
p22_pop	float64	
p23_pop	float64	

Code Référentiel Géographie		
Nom	Type	Description
COM_CODE	string	Code commune
COM_NOM	string	Nom complet
AUC_NOM	string	Nom centre aire urbaine
ATLAS_NOM	string	Nom Atlas
geolocalisation	string	Coordonnées GPS à refaire en longitude et latitude dans BI)

a-com2021-2154 (json)/geopandas	
codegeo	objet
libelle	string
geometry	objet (à refaire en longitude et latitude dans BI)

Dim_entreprises instant_T		
etablissements		
colonnes	types	description
ident	string	Identifiant unique de l'unité légale (9 chiffres)
siren	string	Identifiant unique de l'établissement (SIREN + NIC)
dateCreationEtablissement	date	Date de création de l'établissement (AAAA-MM-JJ)
dateCreationUnitLegale	date	Date de création de l'unité légale
activitePrincipaleUnitLegale	string	Code NAF / APE de l'activité principale de l'unité légale
categorieJuridiqueUnitLegale	string	Catégorie juridique de l'unité légale
trancheEffectifsUnitLegale	string	Tranche d'effectif salarié de l'unité légale (liste de codes)
categorieEntreprise	string	Catégorie d'entreprise (PME, ETI, GE...)
enseigneEtablissement	string	Première ligne d'enseigne commerciale de l'établissement
denominationUsuelleEtablissement	string	Dénomination usuelle de l'établissement
activitePrincipaleEtablissement	string	Code NAF de l'activité principale de l'établissement

Dim_etablissement_principal		
etablissements		
colonnes	types	description
numeroVoieEtablissement	string	Numéro de voie de l'adresse de l'établissement (dim_adresse)
libelleVoieEtablissement	string	Nom de la voie de l'établissement (dim_adresse)
typeVoieEtablissement	string	Type de voie (rue, avenue, boulevard...) (dim_adresse)
codePostalEtablissement	string	Code postal de l'adresse de l'établissement (dim_adresse)
libelleCommuneEtablissement	string	Nom de la commune de l'établissement (dim_adresse)
codeCommuneEtablissement	string	Code INSEE de la commune de l'établissement (dim_adresse)
identifiantAdresseEtablissement	string	Identifiant unique de l'adresse, ou [ND] en diffusion partielle (dim_adresse)
coordonneesLambertAbscisseEtablissement	string	Abcisse des coordonnées Lambert de l'adresse (geolocalisation) (dim_adresse)
coordonneesLambertOrdonneeEtablissement	string	Ordonnée des coordonnées Lambert de l'adresse (geolocalisation) (dim_adresse)

Dim_naf_25		
naf_sections		
colonnes	types	description
NAF 2025 sections	string	Code section (niveau le plus large).
libellés	string	Libellé de la section.
naf_divisions		
colonnes	types	description
NAF 2025 divisions	string	Code division (niveau intermédiaire).
libellés	string	Libellé de la division.
naf_groupes		
colonnes	types	description
NAF 2025 groupes	string	Code groupe (niveau fin).
libellés	string	Libellé du groupe.
naf_classes		
colonnes	types	description
NAF 2025 classes	string	Code classe (niveau très fin).
libellés	string	Libellé de la classe.
naf_sous_classes		
colonnes	types	description
NAF 2025 sous-classes	string	Code sous-classe (niveau le plus détaillé).
libellés	string	Libellé de la sous-classe.

Dim_cat_jurique		
cj_niv1		
code	string(1)	
libellé	string	
cj_niv2		
code	string(2)	
libellé	string	
cj_niv3		
code	string(4)	
libellé	string	

Construction du DWH en 3 étapes

3. CORE

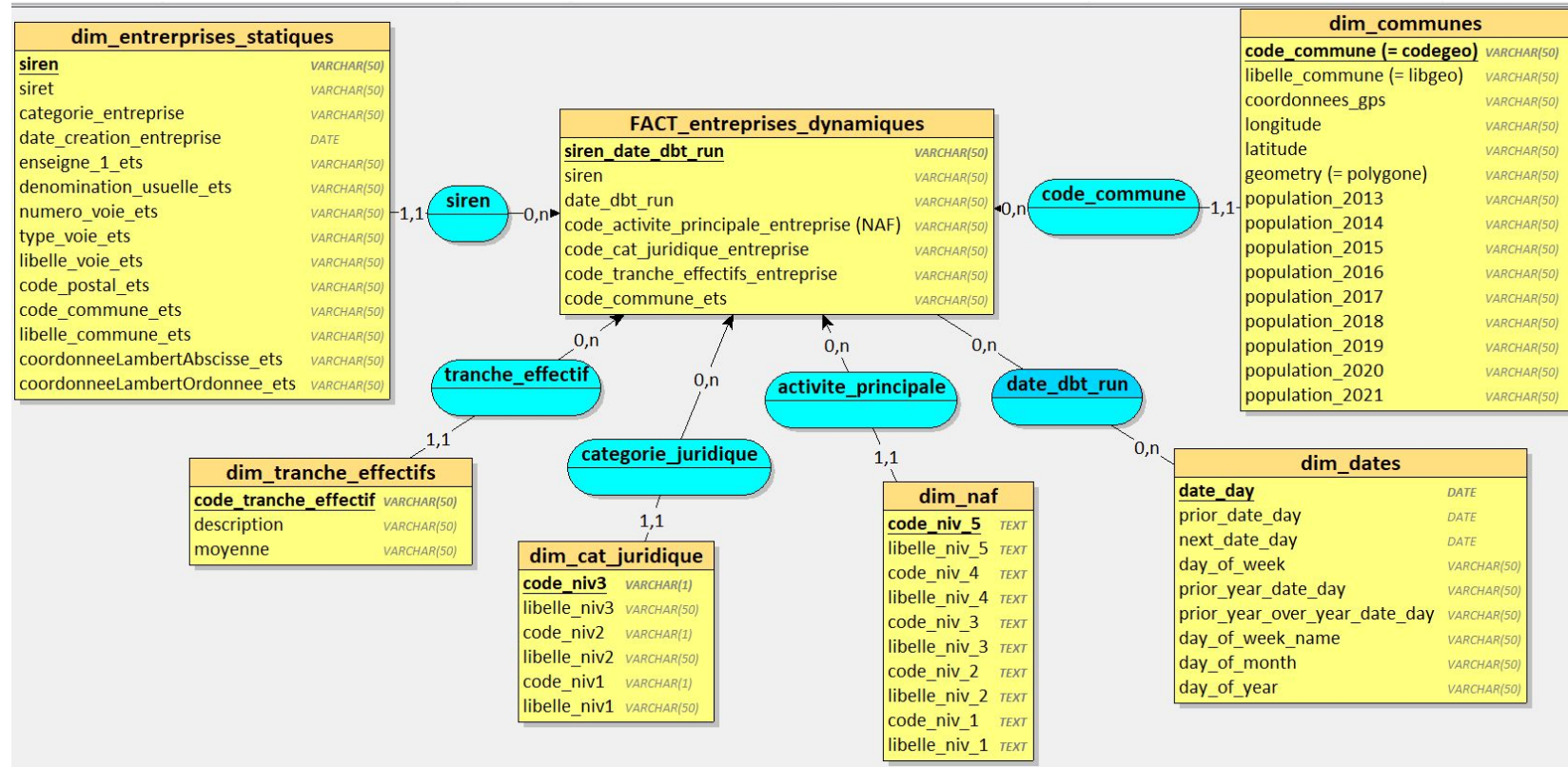
dwh_core	
Tables	
> core_dim_cat_juridique	72K
> core_dim_communes	1,1M
> core_dim_dates	3,6M
> core_dim_effectifs	16K
> core_dim_entreprise_statique	4,2M
> core_dim_entreprises_statiques	5,5M
> core_dim_naf	264K
> core_fact_entreprises_dynamiques	7M

Modèle en étoile :

- **Table de faits** = informations évolutives
- **Dimensions** = informations statiques + précisions en lien avec notre problématique

Expérience :

Structuration en 3 couches claires et efficace, permettant de nombreux allers-retours dans les étapes précédentes pour un rendu final qualitatif
= pipeline clair, reproductible et maintenable



Résultats et enseignements



■ Mise en place d'un datawarehouse conforme aux bonnes pratiques BI



■ Dashboards lisibles, robustes, adaptés à plusieurs publics



■ Développement de compétences avancées:

DBT

Modélisation dimensionnelle

Power BI avancé

Travail collaboratif en mode projet

Conclusion

- Projet complet de BI : de la donnée brute → insight exploitable
- Coopération efficace à 3, structuration & montée en compétence
- DWH pérenne capable d'intégrer de nouvelles données
- Dashboards répondant à des problématiques réelles