

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 ?

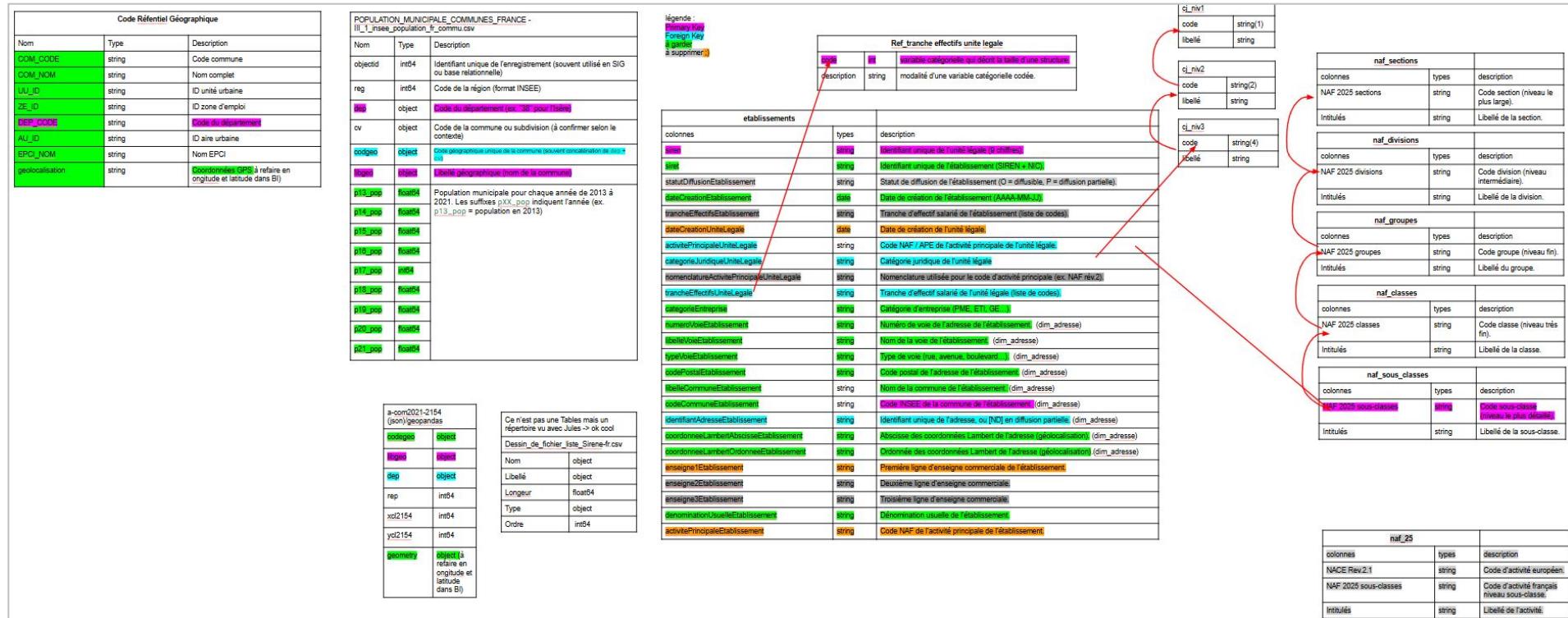
Quelle concentration
d'entreprises par
territoire ?

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

Construction du DWH en 3 étapes

1. RAW

dwh_raw		
		Tables
>	16K	cj_septembre_2022_niveau1
>	16K	cj_septembre_2022_niveau2
>	56K	cj_septembre_2022_niveau3
>	5,5M	dates
>	10M	etablissements
>	46M	geopandas
>	46M	geopandas_someia
>	72K	naf2008_5_niveaux
>	16K	naf2008_liste_n1
>	16K	naf2008_liste_n2
>	56K	naf2008_liste_n3
>	88K	naf2008_liste_n4
>	96K	naf2008_liste_n5
>	3,6M	population_municipale_communes_france
>	680K	ref_geo_clean
>	16K	ref_tranche_effectifs_unite_legale



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



Construction du DWH en 3 étapes

3. CORE

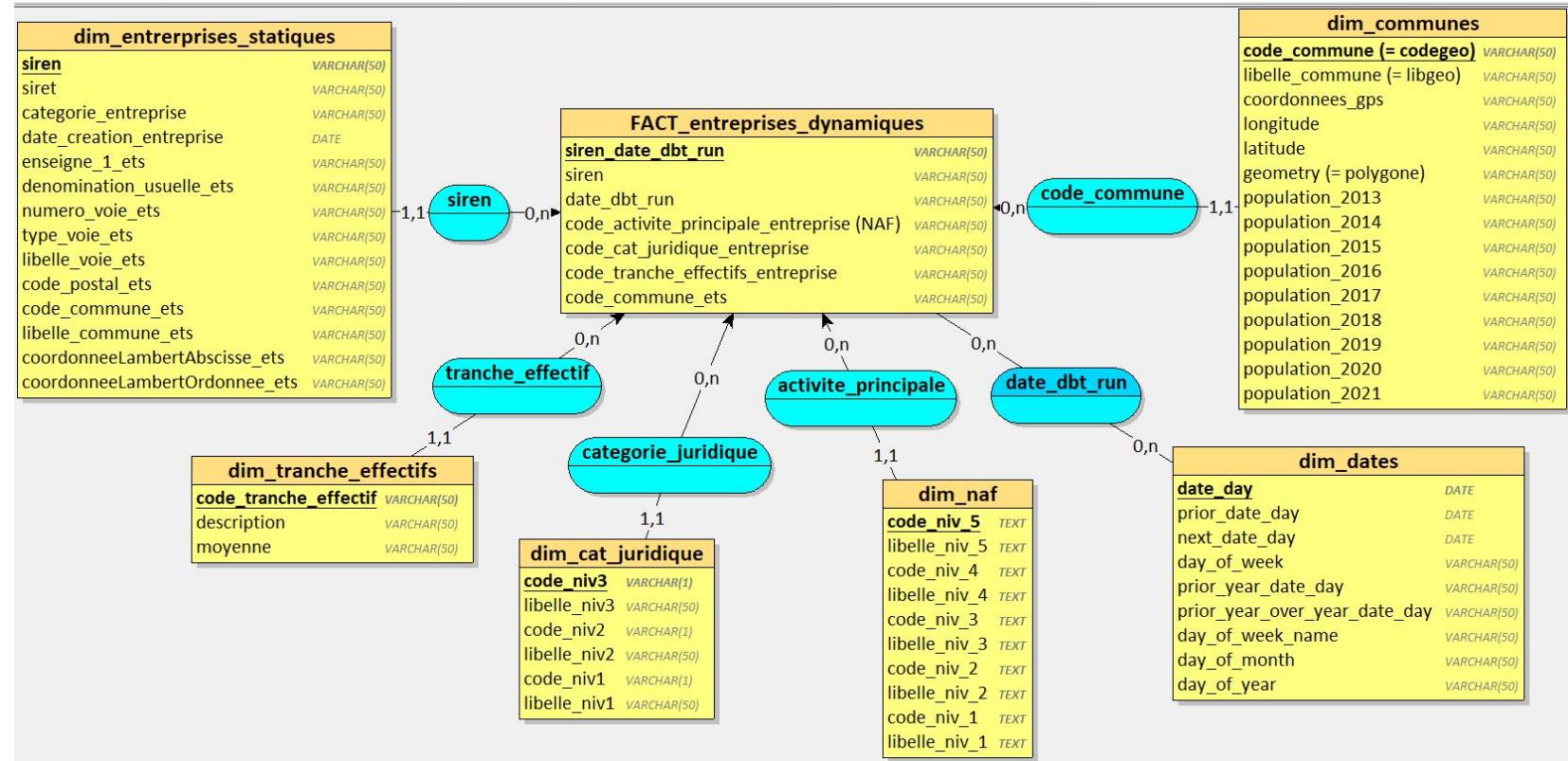
dwh_core	
	Tables
> core_dim_cat_juridique	72K
> core_dim_communes	1,1M
> core_dim_dates	3,6M
> core_dim_effectifs	16K
> core_dim_entreprise_statique	4,2M
> core_dim_entreprises_statiques	5,5M
> core_dim_naf	264K
> core_fact_entreprises_dynamiques	7M

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, réproductible 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