

Accueil

🕒 70 heures

Mis à jour le vendredi 28 juillet 2023

Si vous avez commencé ce projet avant le 15/02/23, vous avez débuté votre travail sur **ce projet archivé**. Vous pouvez continuer l'existant ou recommencer avec ce nouveau projet, dont les modifications s'alignent avec les retours étudiants. Si vous voulez continuer avec la version précédente, veuillez à spécifier "version antérieure au 15 février 2023" dans vos livrables.

Avant de démarrer votre travail sur ce projet, nous vous conseillons de:

- 1 / lire le scénario en entier, chaque section du projet ainsi que les documents fournis ;
- 2 / consulter le **guide des étapes clés** avec des recommandations et des ressources pour organiser votre travail ;
- 3 / préparer une liste de questions pour votre première session de mentorat.



Vous êtes data analyst chez Laplace Immo, un réseau national d'agences immobilières. Le directeur général est sensible depuis quelque temps à l'importance des données, et il pense que l'agence doit se démarquer de la concurrence en créant un modèle pour mieux prévoir le prix de vente des biens immobiliers.



Laplace Immo

Dans ce cadre, elle vous a confié la modification de la base de données permettant de collecter les transactions immobilières et foncières en France. Vous utiliserez ensuite cette base pour analyser le marché et aider les différentes agences régionales à mieux accompagner leurs clients.

À votre arrivée ce matin, vous avez reçu un e-mail de la part de Clara, qui donne plus de détails sur ce qui est attendu de votre part.

Objet : Réunion de validation – Modification de la base de données.

De : Clara

À : Moi

Hello,

Afin d'avancer sur le projet DATAImmo, je prévois une première réunion pour valider la modification de la base de données.

Tu trouveras en pièce jointe un fichier zip avec les données suivantes :

- des données extraites du site open data des Demandes de valeurs foncières (DVF) ;
- des données de l'INSEE avec les résultats des recensements de la population ;
- des données de data.gouv sur les régions, avec le référentiel géographique français, communes, unités urbaines, aires urbaines, départements, académies, régions.

D'ici la réunion, j'aurais besoin de 2 choses de ta part :

1. Il faut que tu prépares le dictionnaire des données en respectant le template en PJ pour répertorier et décrire les données importantes à stocker car nous avons omis de le faire. Il faut le remplir pour les trois fichiers de données.
2. Ensuite, peux-tu modifier le schéma relationnel en pièce-jointe pour qu'il prenne en compte les nouvelles données région et population ? Ça nous permettra de bien visualiser les différentes entités, associations et cardinalités de la base de données. Enfin, il faudra que tu me présentes ce nouveau schéma relationnel normalisé (Il doit suivre la norme 3NF) de la base de données qui donnera lieu à la création des nouvelles tables.

On validera tout ça ensemble lors de notre réunion de validation, pour que tu puisses avancer sur la création de la base.

Je reste à ta disposition en cas de besoin.

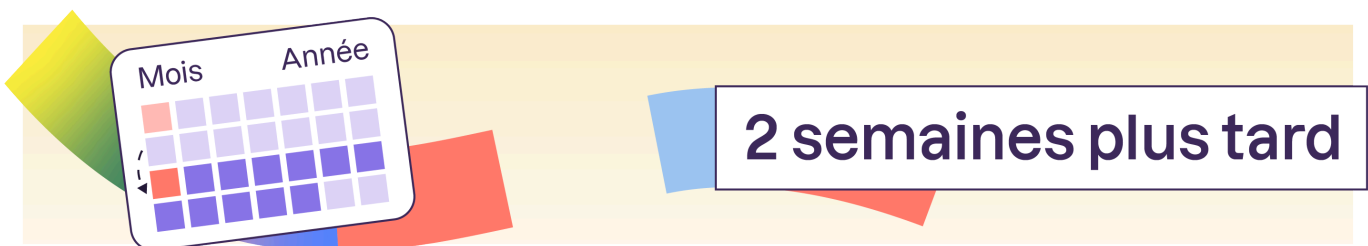
Bien à toi,

Clara Daucourt

CTO

Pièces jointes :

- [Template du dictionnaire des données](#)
- [Les données](#)
- [Le schéma relationnel à recréer et modifier](#)



À ce stade du projet, vous pouvez valider avec votre mentor que votre conception est correcte avant d'aller plus loin. Félicitations ! Elle vous envoie un e-mail à la suite de cette réunion.

Objet: Validation du schéma relationnel et suite du projet

De : Clara

À : Moi

Hello,

Félicitations pour cette première étape de conception ! Comme on en a parlé en réunion ce matin, on est bons pour partir sur la modification de la base de données. Il faut maintenant que tu implémentes les tables dans la base de données en respectant ce qu'on s'est dit ce matin (cf. compte rendu de réunion en pièce jointe).

On a échangé en interne, et on pense qu'un outil comme SQLite est pertinent pour ce type d'implémentation. Si tu es plus à l'aise, tu peux essayer d'implémenter un outil comme MySQL ou PostgreSQL.

Une fois que tu auras fait ça, tu pourras faire les requêtes pour extraire les données dont nous avons besoin. Pour ça, envoie-moi un document PDF avec tous les résultats des requêtes ainsi que la requête associée. Utilise des alias pour que ça soit plus lisible. Tu trouveras toutes les demandes auxquelles il faudra répondre dans le CR de réunion.

Une fois que tu auras terminé, on regardera tout ça ensemble pour voir ce que tu as réussi à tirer comme données.

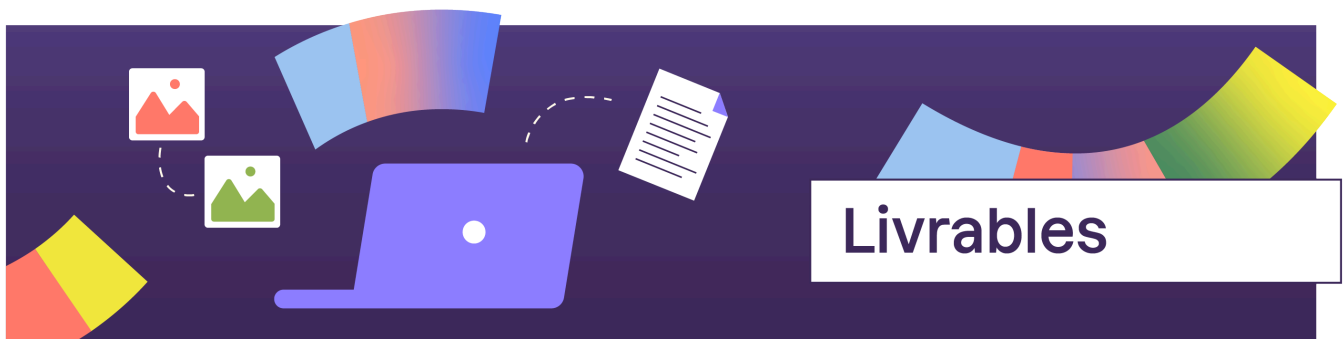
Bon courage !

Clara Daucourt

CTO

Pièce jointe : [Compte-rendu de réunion](#)

Pour vous aider à avancer sur votre projet, nous vous proposons un **guide des étapes clés** sur lequel vous pouvez baser votre travail. Chaque étape propose des recommandations à garder en tête, des problèmes à éviter et des ressources à utiliser.

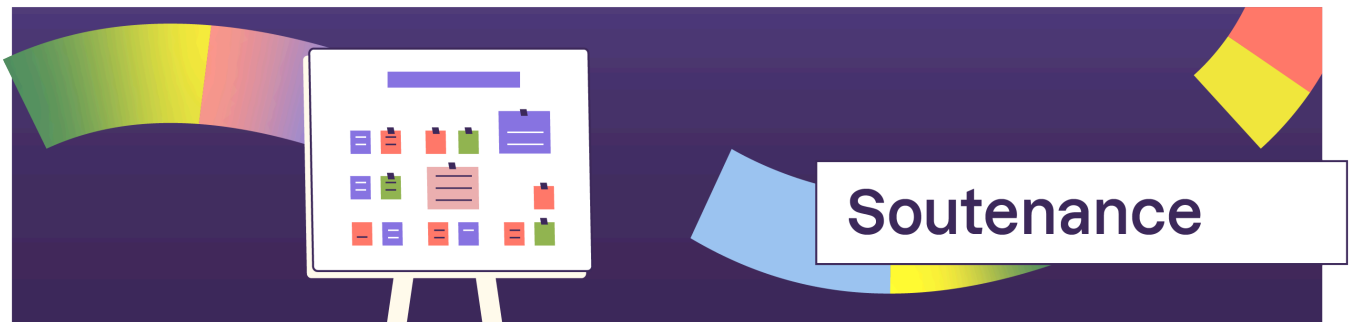


1. Dictionnaire des données.
2. Schéma relationnel normalisé modifié.
3. Fichier zip avec captures d'écran des 4 tables de la base de données opérationnelle, qui incluent le nombre de lignes chargées dans la base :
 - Bien
 - Vente
 - Commune

Pour faciliter votre passage devant le jury, déposez sur la plateforme, dans un dossier zip nommé **"Titre_du_projet_nom_prénom"**, tous les livrables du projet comme suit : **Nom_Prénom_n° du livrable_nom du livrable_date de démarrage du projet**. Cela donnera :

- *Nom_Prénom_1_dictionnaire_de_donnees_mmaaaa*
- *Nom_Prénom_2_schéma_mmaaaa*
- *Nom_Prénom_3_bdd_mmaaaa*
- *Nom_Prénom_4_requetes_mmaaaa*

Par exemple, le premier livrable peut être nommé comme suit : *Dupont_Jean_1_dictionnaire_de_donnees_012023*.



Pendant la soutenance, l'évaluateur jouera le rôle de Clara, la CTO, à qui vous présentez votre travail. Vous pouvez créer un support de présentation mais ce n'est pas obligatoire :

- **Présentation** (15 minutes)
 - Vous expliquerez la démarche de la création de la base de données.
 - Vous présenterez la base de données avec les données importées directement dans l'outil SGBD en live (en réalisant une démonstration, sans vous appuyer sur des captures d'écran).
 - Vous répondrez aux analyses demandées par le Directeur Général.
- **Discussion** (10 minutes)
 - L'évaluateur, jouant le rôle de Clara, vous challengera sur vos choix.
- **Débriefing** (5 minutes)
 - À la fin de la soutenance, l'évaluateur arrêtera de jouer le rôle de Clara pour vous permettre de débriefing ensemble.

Votre présentation devrait durer 15 minutes (+/- 5 minutes). Puisque le respect des durées des présentations est important en milieu professionnel, les présentations en dessous de 10 minutes ou au-dessus de 20 minutes peuvent être refusées.

Compétences évaluées



Charger des données dans une base de données



Créer le schéma d'une base de données



Effectuer des requêtes SQL pour répondre à une problématique métier



Mettre à jour un catalogue de données

Suivant