

Scénario

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.



Ce projet stratégique est appelé en interne le projet « DATAImmo ». La CTO de l'entreprise, Clara Daucourt, a la responsabilité de conduire le projet.

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

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

D'ici là, j'aurais besoin de 3 choses de ta part :

1. Tu trouveras en pièce jointe les données extraites du site open data des [Demandes de valeurs foncières \(DVF\)](#). 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 faudra également rajouter les données de région et de population. Nous pourrons ainsi élargir notre étude. Voici deux idées de sources :

- [Le site de l'INSEE avec les résultats des recensements de la population de l'INSEE](#)
 - [Le site de data.gouv.fr avec le référentiel géographique français, communes, unités urbaines, aires urbaines, départements, académies, régions](#)
2. Ensuite, peux-tu modifier le schéma relationnel 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. Pour le formalisme, nous avons choisi la méthode UML. Tu trouveras le schéma relationnel actuel en pièce jointe.
 3. Enfin, il faudra que tu me présentes ce nouveau schéma relationnel 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](#)



Une fois cette première partie de votre travail effectuée, vous présentez votre travail à Clara qui valide votre schéma relationnel ainsi que le dictionnaire des données.

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).

Une fois que tu auras fait ça, tu pourras faire les requêtes pour extraire les données dont on a 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.

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 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 avancer sur votre projet, nous vous proposons un [guide des étapes clés](#) sur lequel vous pouvez baser votre travail et votre avancement sur le projet. Chaque étape propose des recommandations à garder en tête, des problèmes à éviter et des ressources à utiliser.

Si vous avez commencé ce projet avant le 14/06/22, vous avez débuté votre travail sur **ce projet archivé** et notamment [ce compte rendu de réunion](#) (le lien vers ce document dans le projet archivé a été remplacé par le nouveau compte rendu par erreur). 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 14 juin 2022” dans vos livrables.

Livrables

1. Dictionnaire des données.
2. Schéma relationnel modifié.
3. Base de données opérationnelle avec les données du 1er semestre 2020 (captures d'écran directement du système de gestion de base de données des tables construites).
4. Document avec les requêtes et les résultats (pdf).

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_012022*.

Soutenance

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 débriefer 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

- Créer des tables dans une base de données
- Effectuer des requêtes SQL pour répondre à une problématique métier
- Mettre à jour un catalogue de données
- Charger des données dans une base de données
- Créer le schéma d'une base de données