

Documentation de Projet : DATAImmo

Objectif : Création d'un modèle de prévision des prix immobiliers par l'analyse des données de transactions foncières (DVF) et démographiques (INSEE).

1. Analyse du Contexte et Objectifs

Le projet DATAImmo vise à transformer la masse de données immobilières brutes en un outil stratégique d'aide à la décision.

Missions réalisées :

1. **Structuration** : Création d'un dictionnaire de données complet.
2. **Modélisation** : Conception d'un schéma relationnel normalisé (3NF) intégrant les données géographiques et de population.
3. **Mise en conformité** : Validation du respect du RGPD.
4. **Implémentation** : Création d'une base de données fonctionnelle sous Mysql workbench
5. **Analyse** : Extraction d'indicateurs clés via des requêtes SQL.

2. Conception de la Base de Données

A. Dictionnaire des Données

Les données ont été segmentées en plusieurs tables logiques pour assurer l'intégrité du système :

- **Ventes** : ID transaction, date de mutation, valeur foncière.
- **Biens** : Adresse, surface réelle, nombre de pièces
- **3 tables Géographique** : Communes (avec population INSEE), Départements, Régions.
- **Natures des mutations** : Clé Primaire fictive, Libelle Nature_mutation (Ventes, location)
- **Type de bien** : Clé Primaire fictive, Libelle type_bien (Maison, appartement)

B. Modèle Relationnel (MPD)

Le schéma a été normalisé en **3ème Forme Normale (3NF)** pour éviter toute redondance.

- *Exemple de relation* : Une Vente est liée à un Bien et a une nature de mutation, qui possède un type de bien et qui appartient à une Commune, elle-même rattachée à un Département, puis à une Région.

3. Conformité RGPD et Sécurité

Le projet manipule des données publiques (DVF/INSEE). La conformité a été assurée par :

- **Anonymisation** : Aucune donnée nominative (nom des acheteurs/vendeurs) n'est stockée dans la base.
- **Finalité** : Les données sont utilisées exclusivement pour l'analyse statistique et la prévision du marché immobilier, conformément à la mission de Laplace Immo.

4. Analyse des Données (Requêtes SQL)

Conformément au compte-rendu de réunion, la base a permis d'extraire les indicateurs suivants pour le 1er semestre 2020 :

- Le volume de ventes, via un décompte total des appartements vendus.
- Une analyse régionale grâce à la répartition des ventes par régions.
- L'analyse typologique des biens vendu en requêtant la proportion des ventes selon le nombre de pièces.
- Le Prix au m² et établissement d'un palmarès national des départements les plus chers.
- L'étude de l'évolution des ventes entre le premier trimestre et le second, puis l'évolution sur les 6 premiers mois.
- L'évaluation des performances pour les biens de type T4 et plus, puis tri selon le prix au m² moyen des régions.

5. Livrables Techniques

Les documents suivants (joints à cette documentation) constituent la preuve de l'implémentation :

1. **Dictionnaire_de_donnees_2025.xlsx** : Référentiel technique des champs.
2. **Support_presentation.pdf** : Synthèse visuelle des analyses et graphiques (évolution des ventes, prix moyens).
3. **Requetes.pdf** : Ensembles des requêtes d'analyses demandé pour le projet DataImmo

6. Conclusion et recommandations

- Transformation de données brutes en base fonctionnel et sécurisé.
- Visualisation du marché immobilier 2020 permettant d'orienter la stratégie de l'agence.

Recommandation :

- Affinage des prédictions de prix en intégrant des données socio-économiques.