

Introduction

Contexte

Analyse du marché des assurances habitation en utilisant des données factices pour développer des compétences SQL.

Objectifs principaux

- *Explorer les données fournies.*
- *Construire une méthodologie SQL cohérente.*

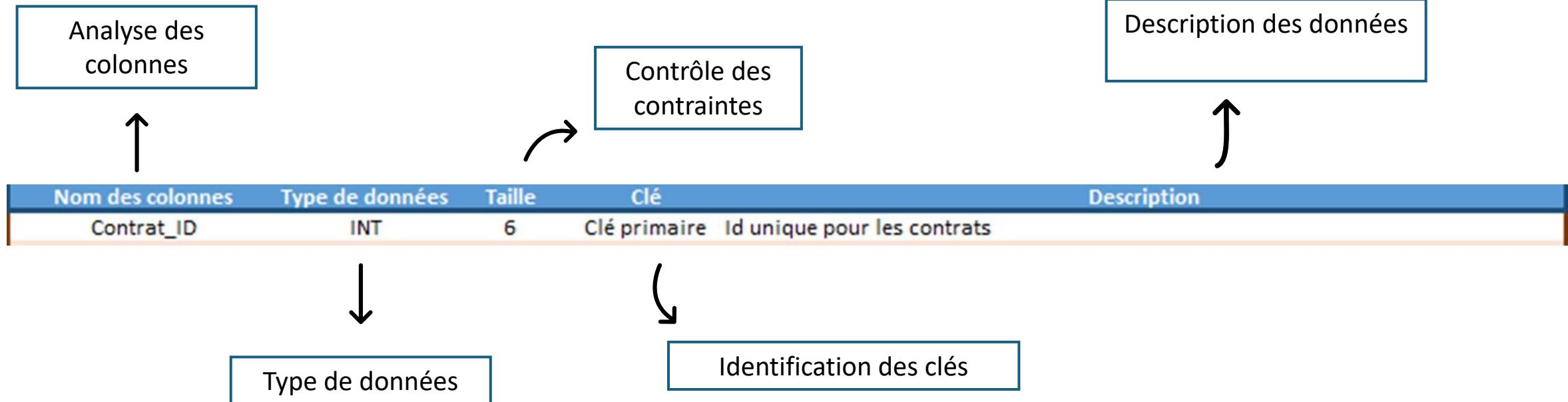
Outils utilisés

Excel (filtres, tableaux, fonctions) et Analyse Exploratoire des données (EDA)

Dictionnaire des données

Objectif principal :

Créer un dictionnaire de données pour documenter les informations clés des fichiers.

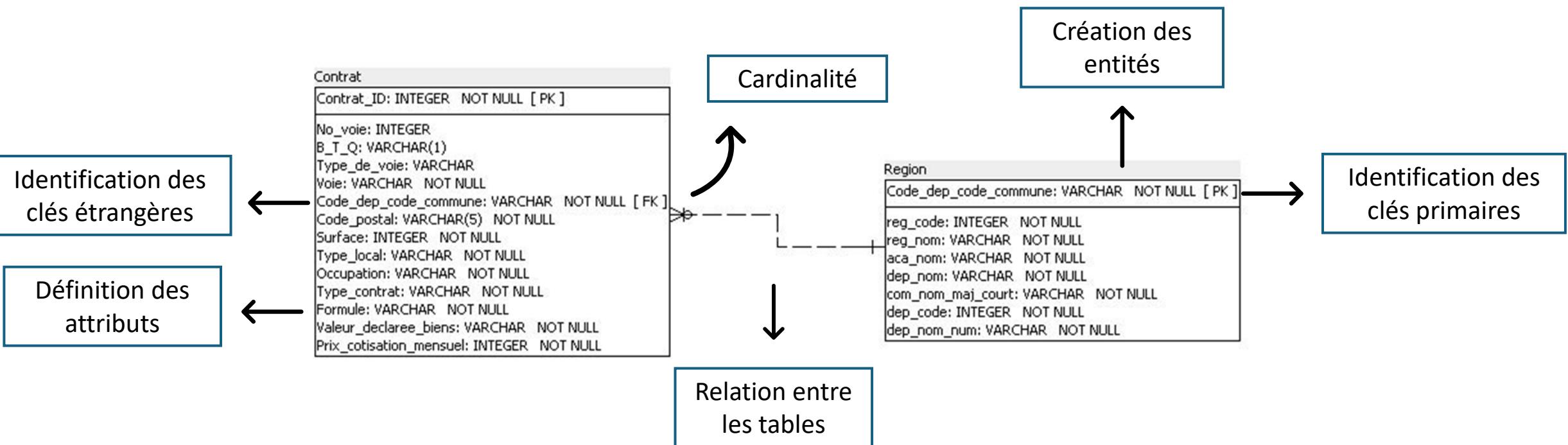


	Nom des colonnes	Type de données	Taille	Clé	Description
CONTRAT.CSV	Contrat_ID	INT	6	Clé primaire	Id unique pour les contrats
	No_voie	INT	4		Numéro dans la voie pour l'adresse du logement assuré
	B_T_Q	CHAR	1		Indicateur éventuel de répétition pour l'adresse du logement assuré sur un caractère
	Type_de_voie	VARCHAR	4		Type de voie pour l'adresse du logement assuré: rue, av (Avenue), rte (Route), ...
	Voie	VARCHAR			Libellé de la voie pour l'adresse du logement assuré
	Code_dep_code_commune	VARCHAR	6	Clé secondaire	Concaténation du code département et code commune pour avoir une clé unique
	Code_postal	INT	5		Code postal pour l'adresse du logement assuré
	Surface	INT	3		Superficie totale du bien assuré, exprimé en mètres carrés
	Type_local	VARCHAR	11		Catégorie du bien immobilier assuré
	Occupation	VARCHAR	12		Statut de l'occupant du bien
	Type_contrat	VARCHAR	19		Nature de l'usage du bien dans le contrat d'assurance
	Formule	VARCHAR	9		Niveau ou catégorie de la couverture d'assurance
	Valeur_declaree_biens	VARCHAR	11		Plage de la valeur déclarée des biens assurés dans le logement
	Prix_cotisation_mensuel	INT	3		Montant de la cotisation d'assurance mensuel exprimée en euro
REGION.CSV	Code_dep_code_commune	VARCHAR	6	Clé primaire	Concaténation du code département et code commune pour avoir une clé unique
	reg_code	INT	2		Code numérique de la région selon la nomenclature officielle
	reg_nom	VARCHAR			Nom complet de la région administrative
	aca_nom	VARCHAR			Nom de l'académie administrative associée à la région
	dep_nom	VARCHAR			Nom du département correspondant
	com_nom_maj_court	VARCHAR			Nom de la commune écrit en majuscule
	dep_code	INT			Code numérique officiel du département
	dep_nom_num	VARCHAR			Nom du département avec le code entre parenthèse

Schéma relationnel

Objectif :

Visualiser la structure des tables et leurs relations dans la base de données.



Création de la base de données

Objectif :

Construire une base de données relationnelle pour analyser les données avec des requêtes SQL.

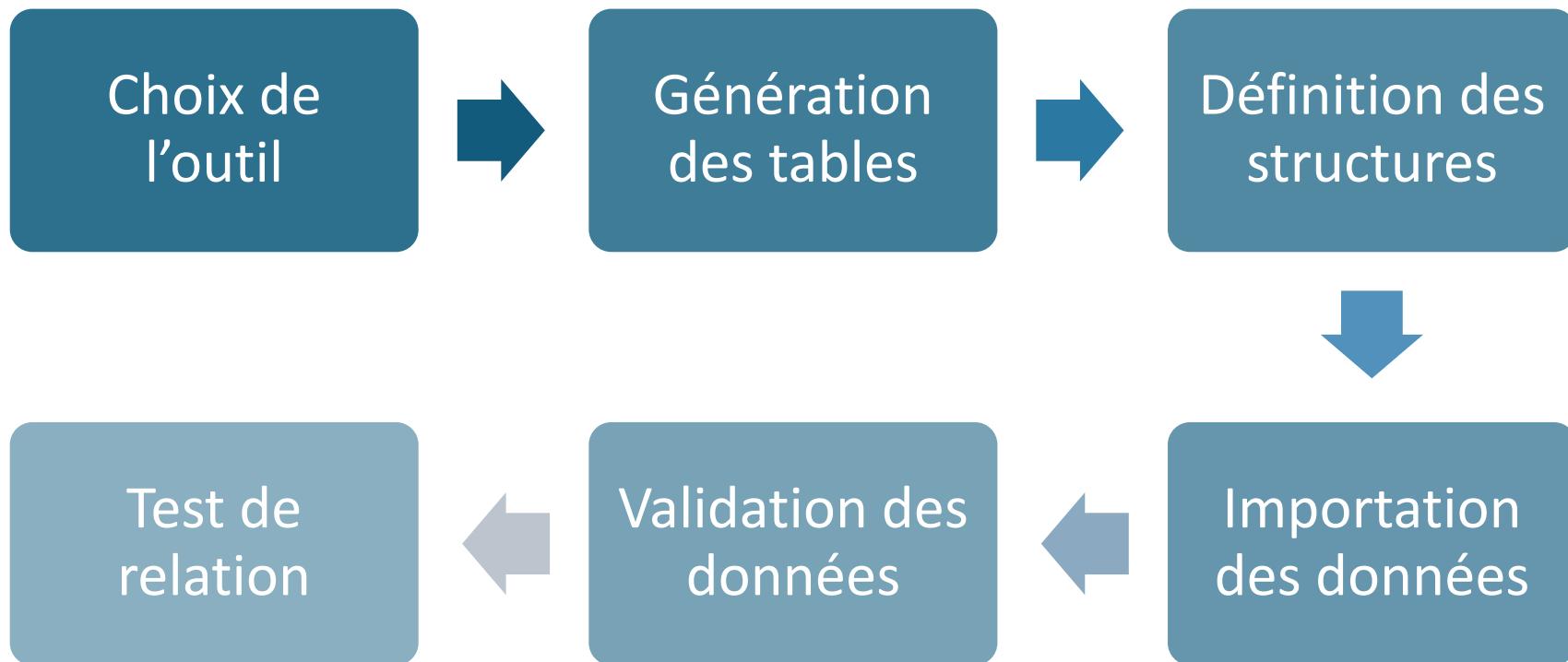


Table Contrat :

The screenshot shows the MySQL Workbench interface for the 'bdd_assurance_habitation' database. The 'Contrat' table is selected. The table has columns: Contrat ID, No voie, B T Q, Type de v, Voie, Code dep, Code pos, Surface, Type local, Occupation, and Type cc. A message at the top right indicates 30335 rows loaded.

Contrat ID	No voie	B T Q	Type de v	Voie	Code dep	Code pos	Surface	Type local	Occupation	Type cc
100001	100	A	TYPE DE V	ROUTE	1200	1270	50	A	Occupation	Type cc

Table Region :

The screenshot shows the MySQL Workbench interface for the 'bdd_assurance_habitation' database. The 'Region' table is selected. The table has columns: Code_dep_, reg_code, reg_nom, aca_nom, dep_nom, com_nom_maj_court, dep_code, and dep_nom_r. A message at the top right indicates 38916 rows loaded.

Code_dep_	reg_code	reg_nom	aca_nom	dep_nom	com_nom_maj_court	dep_code	dep_nom_r
1001	84	Auvergne-Rhône-Alpes	Lyon	Ain	LA BERGERE CLEMENCIAT	1	Ain (01)

Rédiger des requêtes SQL

Objectif :

Construire des requêtes SQL efficaces et adaptées aux besoins d'analyse des données.



Exemple :

Sélectionne la colonne du nom des départements dans la table « Region »

Spécifie la table où les prix de cotisation mensuelle sont enregistrés

Trie par prix moyen décroissant

-- Requête 10 :
-- Classement des 10 départements où le prix moyen de la cotisation est le plus élevé.

SELECT

Region.dep_nom,

ROUND(AVG(prix_cotisation_mensuelle), 2) AS prix_moyen

Alias donné à la colonne calculée

Prix moyen de la cotisation arrondi à deux décimales

FROM

Contrat

JOIN

Region ON Contrat.code_dep_code_commune = Region.code_dep_code_commune

GROUP BY

Region.dep_nom

Regroupe par département

Jointure entre « Contrat » et « Region »

ORDER BY

prix_moyen DESC

LIMIT 10;

Limite le résultat aux 10 premiers départements

Conclusion

Résumé du projet :

Analyser le marché des assurances habitation avec des requêtes SQL.

Compréhension des données

Exploration des fichiers et création d'un dictionnaire de données.

Méthodologie SQL

Utilisation des fonctions SQL pour analyser les données.

Développement des compétences

Maîtrise des requêtes SQL.

Extraction des données pertinentes pour générer des analyses.

Perspectives

Application des compétences SQL dans des projets réels de gestion et analyse de données.