Projet 3: Requêtez une base de données avec SQL



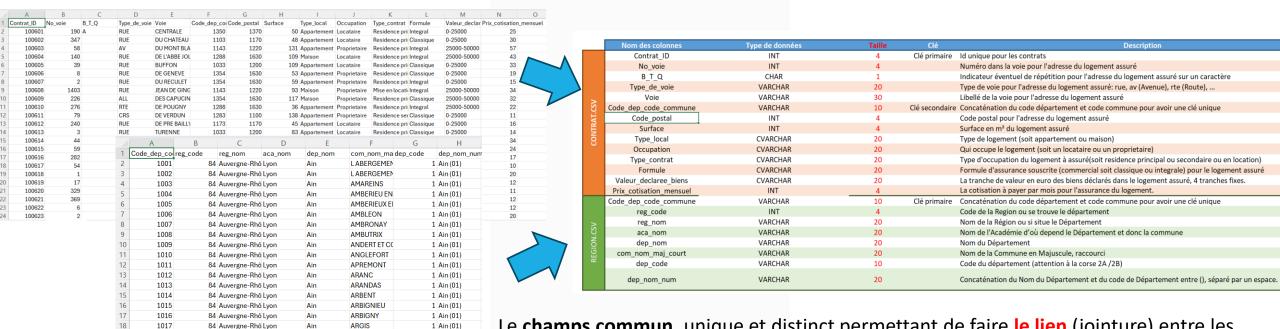
Méthodologie, résultats, et tendances.

# Marché des assurances habitation: objectifs de l'analyse

- L'objectif de la présente étude est de fournir une analyse des données du marché des assurances habitation en France en décrivant les étapes clefs suivantes :
  - Analyse des données brutes provenant des CSV Contrats et Régions.
  - Etablissement du dictionnaire des données.
  - Elaboration du schéma relationnel.
  - Transcription du schéma en Base de Données.
  - Exploitation de la base de données pour l'analyse des tendances par requêtes SQL.
  - Conclusion.

#### Marché des assurances habitation: des données brutes au dictionnaire de données.

L'analyse des données d'entrée des contrats d'assurance croisées avec les informations géographiques des régions françaises permettent d'établir le dictionnaire des données, 1 ère étapes nécessaire à l'élaboration de la base pour stocker ces données clients (BDD)



1017

1018

1019

1020

1021

19

84 Auvergne-Rhô Lyon

84 Auvergne-Rhô Lyon 84 Auvergne-Rhô Lyon **ARGIS** 

ARLOD

ARMIX

ARNANS

ARS SUR FOR

ASNIERES SUF

Ain

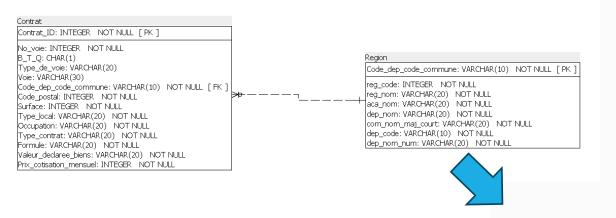
Ain

1 Ain (01)

Le **champs commun**, unique et distinct permettant de faire **le lien** (jointure) entre les contrats et les informations de régions est la concaténation du Code département avec le code commune (colonne Code dep code commune).

#### Marché des assurances habitation: du schéma Relationnel à la création de la base de données

1 même Région identifiée par son Code dep code commune (clef primaire) pouvant concerné plusieurs Contrat d'assurance (identifié par sa Contrat ID, Clef primaire ), la définition des tables avec le type de chaque attribut ainsi que la relation entre les contrats et les Région (via la contrainte de clef étrangère Code\_dep\_code\_commune de la table contrat) peuvent être modélisés afin d'établir le schéma relationnel d'entrée nécessaire à la création de la base de données finale.

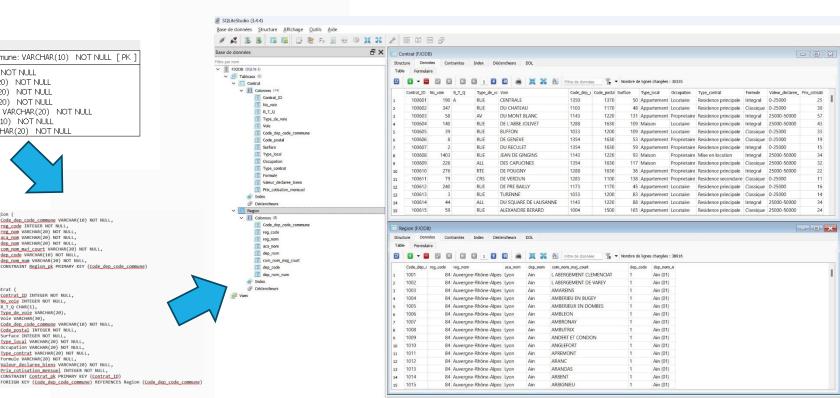


reg code INTEGER NOT NULL, eg nom VARCHAR(20) NOT NULL

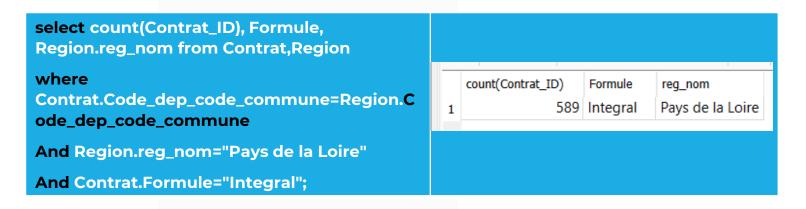
dep code VARCHAR(10) NOT NULL,

/aleur declaree biens VARCHAR(20) NOT NULL, rix cotisation mensuel INTEGER NOT NULL.

La transcription du schéma relationnel en Base de données est établie via les commandes SQL « create table » afin d'être " chargée dans le SGBD avec ensuite toutes les données d'entrée correspondantes de contrats et d'info géographique.



- La base de données étant établie, nous pouvons maintenant interroger, récupérer, et croiser les données nécessaires à l'analyse des marchés d'assurances habitation!
- Nous pouvons filtrer les données souhaitées et ainsi nous concentrer sur une Region particulière par example les Pays de la Loire, afin de croiser les données des contrats qui ont une formule de type « Intégrale » grâce à la relation établie entre les 2 tables suivant le bon Code\_dep\_code\_commune commun correspondant:



 Cette analyse nous permet ainsi de déduire le nombre de contrat Intégral pour cette région particulière des Pays de la Loire s'établissant à 598 contrats!

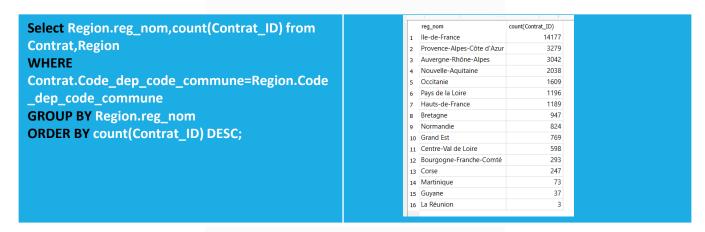
 Nous pouvons maintenant interroger, récupérer, et croiser les données nécessaires à l'analyse des marchés d'assurances habitation dans chaque Régions française.

La table **Régions** nous permet d'avoir une vue d'ensemble **exhaustive de toutes les Régions** de **France** en sélectionnant **toutes** les Régions

distinctes de la table :



Ainsi nous pouvons comptabiliser le nombre de contrat par Région de France en croisant les données avec la table définissant les Contrats grâce à la relation établie entre les 2 tables suivant le bon Code\_dep\_code\_commune commun correspondant, en regroupant au préalable par chaque Nom de Region et en comptant ensuite le nombre de contrat pour chacune d'entre elles, en présentant l'information finale classée par ordre décroissant:



 On distingue clairement et largement la Région « Ile de France » en tête du classement présentant le plus de contrat d'assurances parmi toutes les régions de France!

En analysant plus en détail, cette fois ci à l'échelle des Communes et ayant plus de 150 contrats, toujours en croisant les 2 tables en regroupant au préalable par chaque Nom de Commune et en restreignant le nombre de contrat supérieur ou égale à 150 pour chacune d'entre elles, en présentant l'information finale classée par ordre décroissant:



- On distingue plus précisément les 4 communes (ou plutôt arrondissements) de la Région « Ile de France » vu précédemment que sont Paris 18-ème/
   Paris 17-ème / Paris 15-ème / Paris 16-ème en tête du classement du plus grand nombre de contrat d'assurances!
- Suivant le même concept d'analyse, le prix moyen ou la cotisation est le plus élevé se trouve également dans ce département d'Ile de France: Paris / Hauts de Seine / Val-de-Marnes / Yvelines

Select Region.dep_nom,AVG(Prix_cotisation_mensuel)
from Contrat,Region
WHERE
Contrat.Code_dep_code_commune=Region.Code_dep_
code_commune
GROUP BY Region.dep_nom
ORDER BY AVG(Prix_cotisation_mensuel) DESC
LIMIT 10;

	dep_nom	AVG(Prix_cotisation_mensuel)
1	Paris	36.39908732628086
2	Hauts-de-Seine	26.26935840707965
3	Val-de-Marne	19.82416396979504
4	Yvelines	18.88806970509383
5	Rhône	18.48551829268293
6	Ain	18.23888888888889
7	Alpes-Maritimes	18.13762287756926
8	La Réunion	17.33333333333333
9	Charente-Maritime	17.31730769230769
10	Haute-Savoie	17.14873417721519

 Si l'on compare nos résultats précédents de cotisation les plus élevés avec la moyenne National des cotisations mensuelles tous contrats confondus, on constate que les 3 départements d'Ile De France (Paris / Hauts de Seine / Val-de-Marnes) sont bien au-dessus de cette moyenne Nationale s 'établissant autour des 19 Euros

select AVG(Contrat.Prix\_cotisation\_mensuel)
from Contrat;

AVG(Contrat.Prix\_cotisation\_mensuel)
1 19.33066191748293

En continuant notre analyse sur le département de Paris, nous pouvons estimer également que la surface moyenne des contrats à Paris est d'environs
 52 m²

select AVG(Surface) from Contrat,Region
where
Contrat.Code\_dep\_code\_commune=Region.Code\_dep\_
code\_commune and Region.dep\_nom="Paris";

AVG(Surface)
1 51.76954988591579

## Marché des assurances habitation: conclusion

- Après avoir intégré dans notre base de données les données brutes d'entrée csv des contrats d'assurance et des régions de France, nous pouvons exploitées ces données et croisées ces informations de « Contrat – Region » pour définir les principales tendances en France :
  - la Région « Ile de France » arrive en tête du classement présentant le plus de contrat d'assurances parmi toutes les régions de France
  - Plus précisément 4 arrondissements de la Région « Ile de France » Paris 18-ème/ Paris 17-ème / Paris 15-ème / Paris 16-ème arrive en tête du classement du plus grand nombre de contrat d'assurances
  - Les 3 départements d'Ile De France (Paris / Hauts de Seine / Val-de-Marnes) sont bien au-dessus de la moyenne Nationale des cotisations mensuelles s'établissant à 19 Euros
  - La surface moyenne des contrats à Paris est d'environs 52 m²
  - => Dans la région lle de France et ses départements se concentrent donc les marchés les plus significatifs et importants dans le domaine de l'assurance habitation!

