

Analyse pour Dat 'Assur Habitation

Méthodologie employée

Leonardo MONTILLA

Openclassrooms – BI Analyst – Octobre 2023

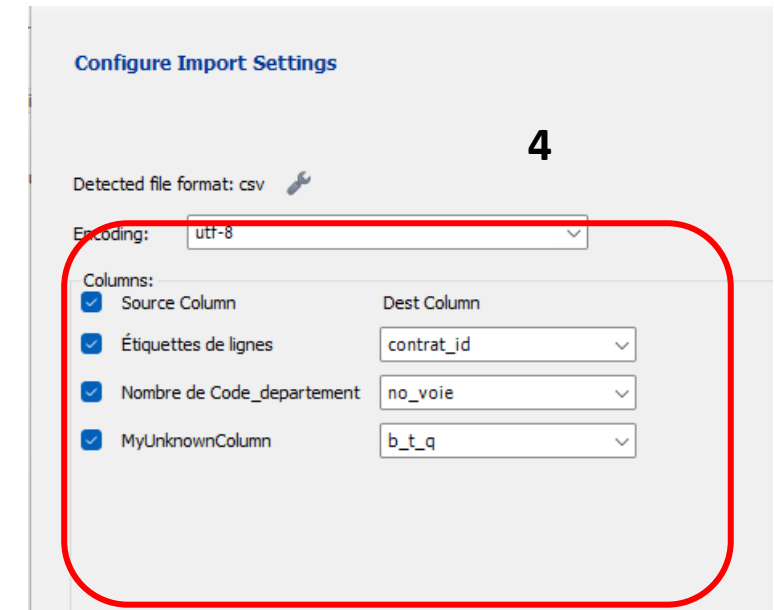
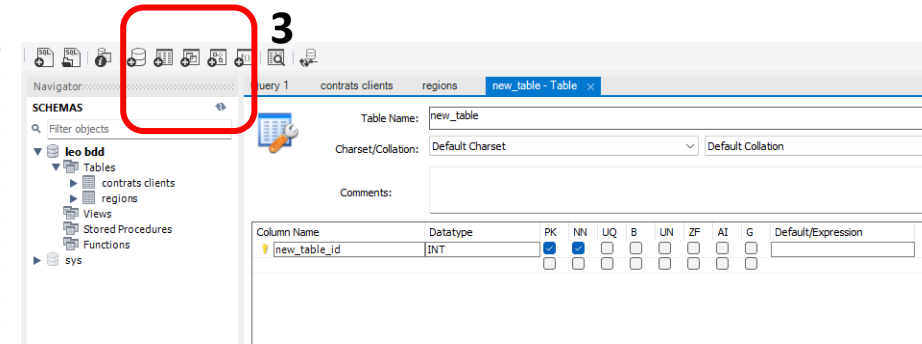


La mise en place des bases de données

code_pos	reg_code	reg_nom	aca_nom	dep_nom	com_nom_maj	dep_code	dep_nom	im
69001	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	AFFOUX	69	hone (69)	
69002	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	AIGUEPERSE	69	hone (69)	
69003	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	ALBIGNY-SUR-SAONE	69	hone (69)	
69004	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	ALIX	69	hone (69)	
69005	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	AMBERIEUX	69	hone (69)	
69006	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	AMPLEPUIS	69	hone (69)	
69007	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	AMPUIS	69	hone (69)	
69008	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	ANCY	69	hone (69)	
69009	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	ANSE	69	hone (69)	
69010	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	L'ARBRESLE	69	hone (69)	
69011	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	ARBUISSONNAS	69	hone (69)	
69012	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	LES ARDILLATS	69	hone (69)	
69013	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	ARNAS	69	hone (69)	
69014	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	AVEIZE	69	hone (69)	
69015	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	AVENAS	69	hone (69)	
69016	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	AZOLETTE	69	hone (69)	
69017	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	BAGNOLS	69	hone (69)	
69018	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	BEAUJEU	69	hone (69)	
69019	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	BELLEVILLE	69	hone (69)	
69020	84	Auvergne-Rhône-Alpes	Auvergne-Rhône-Alpes	Rhône	BELMONT-D'AZERGUES	69	hone (69)	

1

2



- 1) Caractères spéciaux
- 2) Potentielles clés étrangères
- 3) Créer une table
- 4) Correspondance des colonnes

Typologies des données

Types de données

- *VARCHAR (#)*
- *INT*
- *FLOAT*
- *DATE*
- *TIME*
- *BOOLEAN*
- *TEXT*
- ...

Contrats clients

Column	Type	D	Nullable
◇ contrat_id	int		NO
◇ no_voie	int		YES
◇ b_t_q	varchar(10)		YES
◇ type_de_voie	varchar(20)		YES
◇ voie	varchar(50)		YES
◇ code_postal	int		YES
◇ commune	varchar(40)		YES
◇ dep_code	int		YES
◇ surface	int		YES
◇ type_local	varchar(20)		YES
◇ occupation	varchar(20)		YES
◇ type_contrat	varchar(30)		YES
◇ formule	varchar(15)		YES
◇ valeur_declare_mobil...	varchar(20)		YES
◇ prix_cotisation_mens...	int		YES

Régions

Column	Type	C ▲	Nullable
◇ code_postal	int		NO
◇ reg_code	int		YES
◇ reg_nom	varchar(50)		YES
◇ aca_nom	varchar(60)		YES
◇ dep_nom	varchar(100)		YES
◇ com_nom_maj	varchar(50)		YES
◇ dep_code	int		YES
◇ dep_nom_num	varchar(50)		YES

Quelques notions

Dimension

Indicateur

Agrégateur

Projection

Restriction

```
SELECT dep_code, AVG(prix_cotisation_mensuel) AS prix_moyen_cotisation
FROM 'contrats clients'
GROUP BY dep_code
ORDER BY prix_moyen_cotisation
DESC
LIMIT 10
```

```
SELECT
YEAR(Date) AS Année,
Produit,
SUM(Quantité * PrixUnitaire) AS ChiffreAffaires
FROM Ventes
WHERE Produit = 'ProduitX'
AND YEAR(Date) = 2023;
```

Les requêtes réalisées

1. Numéros de contrats avec leur surface pour la commune de Caen

```
1 • select contrat_id, surface
2   from `contrats clients`
3  where commune = 'caen'
```

Result Grid		
Filter Rows:		
	contrat_id	surface
▶	103791	35
	103792	99
	103793	40
	103794	20
✱	NULL	NULL

2. Numéros de contrats avec le type de contrat et leur formule pour les maisons du département de la Saône-et-Loire

```
1 • select distinct c.contrat_id as "contrats maisons", c.type_contrat, c.formule, c.dep_code, r.dep_nom
2   from `contrats clients` as c
3  join regions r on c.dep_code = r.dep_code
4  where c.type_local = 'maison' and c.dep_code = '71'
```

Result Grid					
Filter Rows:					
Export: Wrap Cell Content:					
	contrats maisons	type_contrat	formule	dep_code	dep_nom
▶	114812	Residence principale	Integral	71	Saone-et-Loire
	114782	Residence principale	Classique	71	Saone-et-Loire
	114779	Residence principale	Classique	71	Saone-et-Loire
	114768	Residence principale	Integral	71	Saone-et-Loire

Les requêtes réalisées

3. Nom des régions de France

```
1 SELECT DISTINCT reg_nom AS "nom des regions de france"
2 FROM regions
```

Result Grid	Filter Rows:	Export:	Wrap Cell Content:
nom des regions de france			
▶ Auvergne-Rhone-Alpes			
Hauts-de-France			
Provence-Alpes-Cote d'Azur			
Grand Est			
Occitanie			
Normandie			
Nouvelle-Aquitaine			
Centre-Val de Loire			
Bourgogne-Franche-Comte			
Bretagne			
Pays de la Loire			
Ile-de-France			
Guadeloupe			
Martinique			
Guyane			
La Reunion			
Collectivites d'outre-mer			
Mayotte			

4. Nombre de contrats sur les résidences principales

```
1 • SELECT COUNT(*) AS "nombre de contrats residence principale"
2 FROM `contrats clients`
3 WHERE type_contrat = 'residence principale'
```

Result Grid	Filter Rows:	Export:	Wrap Cell Content:
nombre de contrats residence principale			
▶ 25306			

Les requêtes réalisées

5. Surface moyenne des logements avec un contrat à Paris

```
1 • SELECT AVG (surface) AS "surface_moyenne"  
2 FROM `contrats clients`  
3 WHERE commune LIKE "PARIS%"
```

Result Grid	Filter Rows:	Export:
surface_moyenne		
▶ 51.7695		

6. Les 5 contrats qui ont les surfaces les plus élevées

```
1 • SELECT c.contrat_id, c.surface  
2 FROM `contrats clients` as C  
3 order by c.surface DESC  
4 limit 5
```

Result Grid	Filter Rows:	Edit:
contrat_id	surface	
▶ 104211	815	
105463	742	
130878	595	
100822	570	
109872	559	
* NULL	NULL	

7. Prix moyen de la cotisation mensuelle

```
1 • select avg(prix_cotisation_mensuel) AS 'prix moyen de cotisation mensuel'  
2 from `contrats clients`
```

Result Grid	Filter Rows:	Export:	Wrap Cell Content:
prix moyen de cotisation mensuel			
▶ 19.3720			

Les requêtes réalisées

8. Nombre de contrats pour chaque catégorie de prix de la valeur déclarée en biens

```
1 • SELECT valeur_declare_mobilier, COUNT(*) AS nombre_de_contrats
2 FROM `contrats_clients`
3 GROUP BY valeur_declare_mobilier
4
```

Result Grid		
Filter Rows:		
Export: Wrap Cell Content:		
valeur_declare_mobilier	nombre_de_contrats	
0-25000	22405	
25000-50000	6772	
100000+	104	
50000-100000	695	

9. Classement des 10 départements où le prix moyen de la cotisation est le plus élevé

```
1 • SELECT dep_code, AVG(prix_cotisation_mensuel) AS prix_moyen_cotisation
2 FROM `contrats_clients`
3 GROUP BY dep_code
4 ORDER BY prix_moyen_cotisation
5 DESC
6 LIMIT 10
7
```

Result Grid		
Filter Rows:		
Export: Wrap Cell Content:		
dep_code	prix_moyen_cotisation	
75	36.3991	
92	26.2694	
94	19.8284	
78	18.8779	
69	18.4625	
1	18.2389	
6	18.1422	
17	17.3173	
74	17.1313	
93	16.7100	

10. Nombre de contrats avec des formules « intégral » pour la région Pays de la Loire

```
1 • SELECT COUNT(contrat_id) AS nombre_de_contrats_integral, r.reg_nom AS nom_de_region
2 FROM `contrats_clients` AS c
3 JOIN regions AS r ON c.code_postal=r.code_postal
4 WHERE c.formule = 'integral' AND r.reg_nom = 'Pays de la Loire';
5
```

Result Grid		
Filter Rows:		
Export: Wrap Cell Content:		
nombre_de_contrats_integral	nom_de_region	
587	Pays de la Loire	

Les requêtes réalisées

```
1 • SELECT `commune`, COUNT(*) AS nombre_de_contrats
2 FROM `contrats clients`
3 GROUP BY `commune`
4 HAVING COUNT(*) >= 150
5 ORDER BY `nombre_de_contrats`
6 DESC
```

11. Liste des communes ayant eu au moins 150 contrats

Result Grid			Filter Rows:	Export:	Wrap Cell Cont
	commune	nombre_de_contrats			
►	PARIS 18	515			
	PARIS 17	468			
	PARIS 15	407			
	PARIS 16	394			
	NICE	386			
	PARIS 11	381			
	BORDEAUX	302			
	PARIS 20	302			
	NANTES	291			
	PARIS 19	266			
	PARIS 10	263			
	PARIS 12	252			
	PARIS 14	222			
	GRENOBLE	220			
	PARIS 09	204			
	TOULOUSE	187			
	TOULON	170			
	COURBEV...	163			
	LILLE	161			
	PARIS 03	159			

12. Nombre de contrats pour chaque région

```
1 • SELECT COUNT(distinct contrat_id) AS nombre_de_contrats, r.reg_nom as nom_region
2 FROM `contrats clients` AS c
3 JOIN regions AS r ON c.dep_code = r.dep_code
4 GROUP BY r.reg_nom
5 ORDER BY r.reg_nom
```

Result Grid			Filter Rows:	Export:	Wrap Cell Content: F
	nombre_de_contrats	nom_region			
►	3049	Auvergne-Rhone-Alpes			
	283	Bourgogne-Franche-Comte			
	944	Bretagne			
	598	Centre-Val de Loire			
	764	Grand Est			
	34	Guyane			
	1176	Hauts-de-France			
	14162	Ile-de-France			
	17	La Reunion			
	73	Martinique			
	816	Normandie			
	2029	Nouvelle-Aquitaine			
	1580	Occitanie			
	1192	Pays de la Loire			
	3259	Provence-Alpes-Cote d'Azur			

Des questions ?

Merci pour votre attention

