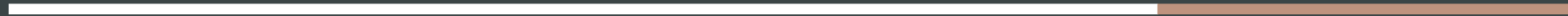


RENDU DE PROJET

Module d'enrichissement géomarketing



1. INTRODUCTION ET CONTEXTE

Objectif du projet : Développer un module d'enrichissement geomarketing

Environnement technique : Python 3.12, modules de géocodage et enrichissement développés avec SQLAlchemy pour la gestion des connexions à la base de données.

MÉTHODOLOGIE

Étapes principales du module :

- Géocodage
- Enrichissement de l'âge
- Enrichissement géomarketing
- Profils de population

Structuration des données :

- Ajout de tables RefCP et Reflris pour le géocodage

GÉOCODAGE

Objectif : Obtenir les codes INSEE et IRIS pour chaque adresse.

Paramètres de la fonction :

- Variables obligatoires (table_entree, top_tnp, cp, ville, id_client).
- Gestion de top_tnp pour distinguer la structure des noms et prénoms (champs combinés ou séparés).

Processus :

- Normalisation des données (normalize_column_name), utilisation de unicode pour un traitement des caractères spéciaux.
- Fusion des tables de référence RefCP et RefIris pour ajouter c_insee et c_iris (et donc codegeo).

GÉOCODAGE

prenom	nom	lieu_dit	cp	ville	id_client	code_postal	code_commune_insee	nom_de_la_commune	c_insee	depcom	lib_iris	code_iris	c_iris	c_qualite_iris	codgeo
Kilian	Chappelle	Longeray-Gare	1500	Ambérieu-en-Bugey	1	1500	1004	AMBERIEU EN BUGEY	01004	01004	Longeray-Gare	010040102	0102	1	010040102
Benjamin	Godin	Les Pérouses-Triangle d'Activités	1500	Ambérieu-en-Bugey	2	1500	1004	AMBERIEU EN BUGEY	01004	01004	Les Pérouses-Triangle d'Activités	010040101	0101	1	010040101
Léa	Durand	Quartier Bellevue	1250	Ceyzériat	3	1250	1072	CEYZERIAT	01072	NaN	NaN	NaN	0000	2	010720000
Léon	Durand	Quartier Bellevue	1250	Ceyzériat	4	1250	1072	CEYZERIAT	01072	NaN	NaN	NaN	0000	2	010720000
Lucas	Durand	Quartier Bellevue	1250	Ceyzériat	5	1250	1072	CEYZERIAT	01072	NaN	NaN	NaN	0000	2	010720000
Olivier	Lefevre	Centre	1220	Divonne-les-Bains	6	1220	1143	DIVONNE LES BAINS	01143	01143	Centre	011430101	0101	1	011430101
Étienne	Dubois	Nord-Ouest	1220	Divonne-les-Bains	7	1220	1143	DIVONNE LES BAINS	01143	01143	Nord-Ouest	011430102	0102	1	011430102

ENRICHISSEMENT DE L'ÂGE :

Objectif : Estimer l'âge et le sexe en utilisant le prénom et la localisation géographique.

Processus :

- Estimation du sexe à partir du prénom.
- Estimation de l'âge par :
 - Géographie (grâce aux données socio-démographiques de tbrefgeo).
 - Prénom (recherche dans la table des prénoms, calcul de l'âge moyen).

ENRICHISSEMENT DE L'ÂGE :

Objectif : Estimer l'âge et le sexe en utilisant le prénom et la localisation géographique.

sexe_civilit	prenom	nom	complement_nom	adresse	complement_adrs	lieu_dit	...	codegeo	c_insee	c_iris	e_sexe	e_age	e_top_age_ok	e_annee_naissance	e_p_5ans	indice_conf_age
M	M	Kilian	Chappelle	None	5 rue des bananes enchantées	None	Longeray-Gare	10040102	10040	0102	M	38.276755	2	1986	0.9	Confiance ++
M	Mme	Benjamin	Godin	None	7 rue michel ange	None	Les Pérouses-Triangle d'Activités	10040102	10040	0102	M	48.600829	2	1975	0.9	Confiance
F	Mme	Léa	Durand	None	2 rue des Charmes	Bâtiment A étage 1	Quartier Bellevue	10720000	10720	0000	F	50.984448	2	1973	0.9	Confiance
M	M	Léon	Durand	None	3 rue des Charmes	Bâtiment A étage 1	Quartier Bellevue	10720000	10720	0000	M	50.984448	2	1973	0.9	Confiance
M	M	Lucas	Durand	None	4 rue des Charmes	Bâtiment A étage 1	Quartier Bellevue	10720000	10720	0000	M	49.880282	2	1974	0.9	Confiance
F	M	Olivier	Lefevre	None	10 rue de la Paix	Appartement 3B	Centre	11430101	11430	0101	F	50.026984	2	1974	0.9	Confiance
None	M	Étienne	Dubois	None	6 allée des Chênes	None	Nord-Ouest	11430102	11430	0102	M	49.210912	2	1975	0.9	Confiance

MERCI
