Sprint qualité du jeu de données des prénoms des nouveaux nés de la ville d'Antibes Juan-les-Pins de 2011 à 2022 par Dieynaba KOUYATE accessible sur https://trouver.datasud.fr/dataset/prenoms-des-nouveaux-nes-de-la-ville-antibes-juan-les-pins

Point de contrôle qualité	Problème observé et commentaire				Correction : comment ? importance ?				
1. le jeu de données est difficilement accessible (format "image", PDF, HTML), le fichier est mal formé	Après vérificat « I ». Nous l'av documentatior adopté le sché bien précisé qu valeurs sont po pour indétermi est télécharge problème d'aff esthétique ma	ion, il y a une o yons détecté gr n via le schéma ema « prénoms ue dans la propossibles (F pou né/interséxué) able au format ichage des car is on comprend	CSV et XLS ma actères accentud d.	omme sexe in a lu la i jeu données a nés »). C'est i _SEXE » 3 ur masculin, I jeu de données	Gravite Fa accentués enregistré apparaisse A 1 COMMUNE 2 Antibes 3 Antibes 4 Antibes 5 Antibes 6 Antibes 7 Antibes 8 Antibes 9 Antibes 10 Antibes 11 Antibes 12 Antibes	ible : Pour i , j'ai visionn sous un for ent bien.	résoudre le né et téléch rmat CSV	e problème des nargé le fichier et les caractère	XLS puis je l'ai
2. la licence est absente ou inhabituelle, le jeu de données n'est pas "open data"	NON Licence ouvert Le « Réutilisat de la reprodui et la transform des produits o la redistribuer, titre commercia	termes de d'Antibes e https://trou	la licence e et en mettar	en donnant nt le lien vo d.fr/datase	omme indiquée t le nom de la d ers les jeux de et/prenoms-des pins	commune données :			

Sprint qualité du jeu de données des prénoms des nouveaux nés de la ville d'Antibes Juan-les-Pins de 2011 à 2022 par Dieynaba KOUYATE accessible sur https://trouver.datasud.fr/dataset/prenoms-des-nouveaux-nes-de-la-ville-antibes-juan-les-pins

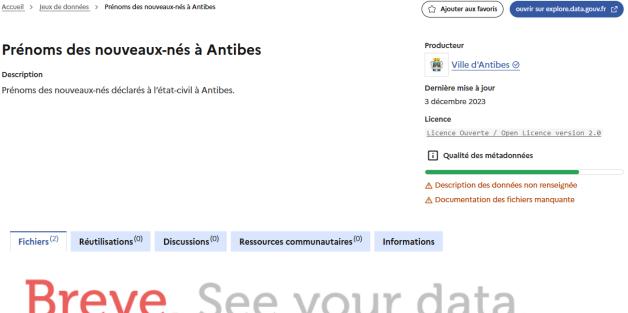
	informations, ou en l'incluant dans son propre produit ou application. Sous réserve de : mentionner la paternité de l' « Information » : sa source (au moins le nom du « Concédant ») et la date de dernière mise à jour de l' « Information » réutilisée. Source https://www.etalab.gouv.fr/licence-ouverte-open-licence/			
3. le fichier fait peu appel aux standards répandu et aux données pivots	NON: Les données respectent le schéma des prénoms https://schema.data.gouv.fr/scdl/prenoms/ . Le code postale indiqué dans le fichier CSV fait bien référence aux Code de l'INSEE: https://www.insee.fr/fr/information/6800675	N.A		
4. le fichier est mal documenté	NON Je n'ai pas de documentation directement sur la plateforme de la ville d'Antibes mais cela s'explique par le fait que ça respecte un schéma et dans le schéma les « Propriétés » sont expliquées ainsi que le caractère obligatoire ou non. On n'a pas directement la documentation sur la plateforme mais le schéma documente bien les champs et attentes. Ce serait bien de l'avoir directement sur le site de la ville d'Antibes dans l'onglet information du jeu de données pour avoir un raccourci et ne pas aller sur schema.data.gouv.fr mais ce n'est pas bloquant car la documentation existe. Il faut aller la chercher.	Gravité faible : Voir le schéma pour comprendre les différents intitulés https://schema.data.gouv.fr/scdl/prenoms/		
OUI D'après l'analyse faite sur https://publier.etalab.studio/fr/ il y a 4 erreurs et on les a toutes trouvées : une colonne renseignée en plus (c'était la Colonne « G » : peut-être une fausse manipulation car ne fait pas partie du schéma), 3 syntaxes de prénoms non valides aux lignes 7310, 7355, 8332.		Gravité Faible : On a supprimé l'élément de la colonne G qui ne fait pas partie du schéma et rectifié la syntaxe des prénoms. Puis on l'a remis sur https://publier.etalab.studio/fr/ et c'était conforme.		

Sprint qualité du jeu de données des prénoms des nouveaux nés de la ville d'Antibes Juan-les-Pins de 2011 à 2022 par Dieynaba KOUYATE accessible sur https://trouver.datasud.fr/dataset/prenoms-des-nouveaux-nes-de-la-ville-antibes-juan-les-pins

Point de contrôle qualité	Problème observé et commentaire	Correction : comment ? importance ?
6. Valeurs aberrantes, suspectes, inexplicables, pas crédibles	NON	N.A
7. Il manque des données et cela n'est pas documenté (trous, données tronquées, valeurs vides, granularité / fréquence / maillage / fraîcheur)	NON	N.A
8. Trop de données : doublons, inutilement vieilles, précision / fréquence / maillage / fraîcheur	NON Il y a des orthographes différentes pour un même prénom mais cela dépend de comment les parents épellent/écrivent le prénom de l'enfant donc ce n'est pas un doublon mais différentes manières d'écrire par exemple « Timothé » et « Timothée » (voir Lignes 2000 et 2001 du fichier CSV)	N.A
9. Données posant problème avec la réglementation (données perso, relatives à la santé, la religion, propriété littéraire et artistique, etc.)	NON Il n'y a que les prénoms, s'il y avait les noms, ce serait une violation des données personnelles mais ce n'est pas le cas.	N.A
10. Les contenus posent problèmes : synonymies, non traduits (USA), cryptique (DAECPP), utilisation du 0 au lieu du "null"	NON	N.A

Sprint qualité du jeu de données des prénoms des nouveaux nés de la ville d'Antibes Juan-les-Pins de 2011 à 2022 par Dieynaba KOUYATE accessible sur https://trouver.datasud.fr/dataset/prenoms-des-nouveaux-nes-de-la-ville-antibes-juan-les-pins

Comme on peut le constater en partie dans l'analyse de Breve : ces données sont de bonnes qualités. D'ailleurs, avant de les choisir nous avons vérifié sur Open data la qualité des métadonnées et c'était vert. Ces données nous semblent tout à fait exploitables et répondent aux critères de l'Open Data.



Breve. See your data.

