

Fichier des prénoms - Édition 2021

lil-1526

La source

Le fichier des prénoms est établi à partir des seuls bulletins de naissance des personnes nées en France (métropole et départements d'outre-mer). En conséquence, l'exhaustivité n'est pas garantie sur toute la période, notamment pour les années antérieures à 1946. Les utilisateurs pourront donc constater des écarts avec le nombre annuel des naissances évalué par l'Insee. Ces écarts, importants en début de période, vont en s'amenuisant. Après 1946, ils sont peu significatifs.

Les informations contenues dans le fichier des prénoms sont basées sur les bulletins d'état civil transmis à l'Insee par les officiers d'état civil des communes. Ces bulletins sont eux-mêmes établis à partir des déclarations des parents. L'Insee ne peut garantir que le fichier des prénoms soit exempt d'omissions ou d'erreurs.

La refonte du processus électoral a entraîné un nombre de corrections dans la base des prénoms plus important que les années précédentes. En effet, chaque électeur est maintenant inscrit au répertoire électoral unique avec son état civil officiel (celui du Répertoire national d'identification des personnes physiques / RNIPP), des anomalies ont donc été corrigées.

Le jeu de données contient deux fichiers

Un fichier national (NAT2021)

Sans localisation géographique de naissance,
Couvrant toutes les années de naissance de 1900 à 2021
Comporte 17 627 prénoms attribués à des personnes de sexe masculin
et 20 854 prénoms attribués à des personnes de sexe féminin.

Un fichier départemental (DPT2021)

Distinguant les départements de naissance,
Couvrant (comme depuis 2014) les années de naissance de 1900 à 2021.
Comporte 17 627 prénoms attribués à des personnes de sexe masculin
et 20 854 prénoms attribués à des personnes de sexe féminin.

Les champs retenus

Les personnes prises en compte

Le champ couvre l'ensemble des personnes nées en France hors Mayotte et enregistrées à l'état civil sur les bulletins de naissance. Les personnes nées à l'étranger sont exclues.

Le champ des prénoms retenus

Dans les fichiers de l'état civil, en l'occurrence les bulletins de naissance, les différents prénoms sont séparés par une espace (ou blanc). Ainsi deux prénoms séparés par un tiret constituent un seul prénom composé (exemple : Anne-Laure). Le premier prénom simple ou composé figure en début de liste, et c'est celui qui sera retenu après le traitement de la protection de l'anonymat.

Conditions portant sur les prénoms retenus

- 1. Sur la période allant de 1900 à 1945, le prénom a été attribué au moins 20 fois à des personnes de sexe féminin et/ou au moins 20 fois à des personnes de sexe masculin
- 2. Sur la période allant de 1946 à 2021, le prénom a été attribué au moins 20 fois à des personnes de sexe féminin et/ou au moins 20 fois à des personnes de sexe masculin
- 3. Pour une année de naissance donnée, le prénom a été attribué au moins 3 fois à des personnes de sexe féminin ou de sexe masculin

Les effectifs des prénoms ne remplissant pas les conditions 1 et 2 sont regroupés (pour chaque sexe et chaque année de naissance) dans un enregistrement dont le champ prénom (PREUSUEL) prend la valeur «_PRENOMS_RARES_». Les effectifs des prénoms remplissant la condition 2 mais pas la condition 3 sont regroupés (pour chaque sexe et chaque prénom) dans un enregistrement dont le champ année de naissance (ANNAIS) prend la valeur «XXXX».

Précision pour le département de naissance

Sur toute la période, le département de naissance (variable DPT) est celui existant au moment de la naissance. Ainsi une personne née à Issy-les-Moulineaux en 1949 sera codée en 75 (Seine), et une personne née en 1971 à Issy-les-Moulineaux sera codée en 92 (Hauts-de-Seine).

En effet, de par la loi n° 64-707 du 10/07/1964, il y a eu création des départements 75 (Paris), 78 (Yvelines), 91 (Essonne), 92 (Hauts-de-Seine), 93 (Seine-Saint-Denis), 94 (Val-de-Marne) et 95 (Val-d'Oise), à partir des anciens départements 75 (Seine) et 78 (Seine-et-Oise), avec une date d'effet dans le fichier au 1^{er} janvier 1968.

Un regroupement de code a été fait pour la variable DPT :

- le code DPT=20 est utilisé pour les naissances ayant eu lieu dans les deux départements 2A (Corse-du-Sud) et 2B (Haute-Corse) ;

Lorsque le champ année de naissance (ANNAIS) est codé «XXXX», le département de naissance (DPT) est également codé en «XX».

Décrets et lois relatifs aux départements métropolitains

- Décret du 18/01/1955 : la Seine-Inférieure devient Seine-Maritime (76)
Décret du 09/03/1957 : la Loire-Inférieure devient Loire-Atlantique (44)
- Loi n° 64-707 du 10/07/1964 : création des départements 75 (Paris), 78 (Yvelines), 91 (Essonne), 92 (Hauts-de-Seine), 93 (Seine-Saint-Denis), 94 (Val-de-Marne) et 95 (Val-d'Oise)

- Décret du 10/10/1969 : les Basses-Pyrénées deviennent Pyrénées-Atlantiques (64)
- Décret du 13/04/1970 : les Basses-Alpes deviennent Alpes-de-Haute-Provence (04)
- Loi n° 75-356 du 15/05/1975 : création des départements 2A (Corse-du-Sud) et 2B (Haute-Corse)
- Décret du 27/02/1990 : les Côtes-du-Nord deviennent Côtes-d'Armor (22)

Éléments de contexte sur le choix des prénoms

Il n'existe pas de liste de prénoms autorisés. Les prénoms connus étrangers ou certains diminutifs peuvent ainsi être choisis. En conséquence, on peut trouver par exemple des prénoms à une seule lettre dans le fichier des prénoms. Ce sont effectivement les parents qui choisissent librement le ou les prénom(s) de l'enfant lors de la déclaration de naissance. En particulier depuis 1993, l'officier d'état civil ne peut plus refuser le prénom choisi par les parents. Il peut toutefois avertir le Procureur de la République s'il estime que le prénom nuit à l'intérêt de l'enfant (exemple : prénom ridicule ou grossier) ; ou que le prénom méconnaît le droit d'un tiers à voir protéger son nom de famille (exemple : un parent ne peut choisir comme prénom le nom de famille d'une autre personne dont l'usage constituerait une usurpation). Le Procureur peut ensuite saisir le Juge aux affaires familiales qui peut demander la suppression du prénom sur les registres de l'état civil. Dans ce cas le prénom serait également supprimé du fichier des prénoms.

L'alphabet utilisé pour l'écriture des prénoms doit être celui qui sert à l'écriture du français. Les caractères alphabétiques qui ne sont pas utilisés dans la langue française ne sont donc pas autorisés (par exemple le « ñ » ou le « ş »). Ainsi, par exemple, le prénom d'origine roumaine "Rareş" apparaît en tant que "RARES" dans le fichier des prénoms.

Description des 2 fichiers

Le corpus est composé de 2 tables :

Tables	Nbre d'observations	Nbre de variables
NAT2021	686 538	4
DPT2021	3 784 673	5

Les variables :

Variables	Attributs	Signification	Table		Valeurs masquées
			Nat	Dpt	
PREUSUEL	Texte 25	Prénom usuel	X	X	«_PRENOMS_RARES_»
SEXE	Texte 1	Sexe	X	X	
ANNAIS	Texte 4	Année de naissance	X	X	«XXXX»
DPT	Texte 3	Département de naissance		X	«XX»
NOMBRE	Numérique	Nombre de naissances	X	X	

Les départements

En plus de 100 ans, les noms des départements et le découpage en départements ont changé. C'est ce dernier point qui est le plus lourd de conséquences.

Sur toute la période, le département de naissance (variable DPT) est celui existant au moment de la naissance. Ainsi une personne née à Issy-les-Moulineaux en 1949 sera codée en 75 Seine, et une personne née en 1971 à Issy-les-Moulineaux sera codée en 92 Hauts-de-Seine.

En effet, de par la loi n° 64-707 du 10/07/1964, il y a eu création des départements 75 (Paris), 78 (Yvelines), 91 (Essonne), 92 (Hauts-de-Seine), 93 (Seine-Saint-Denis), 94 (Val-de-Marne) et 95 (Val-d'Oise), à partir des anciens départements 75 (Seine) et 78 (Seine-et-Oise), avec une date d'effet dans le fichier au 1er janvier 1968.

Par ailleurs, les deux départements 2A Corse-du-Sud et 2B Haute-Corse sont regroupés (DPT=20, Corse), et il en est de même des départements d'Outre-mer qui sont regroupés (DPT=97). En 2013 cela concerne les naissances de la Guadeloupe, de la Martinique, de la Guyane, de La Réunion et de Mayotte (devenue un DOM en 2011).

Décrets et lois relatifs aux départements métropolitains

- décret du 18/01/1955 : la Seine-Inférieure devient Seine-Maritime
- décret du 09/03/1957 : la Loire-Inférieure devient Loire-Atlantique
- loi n° 64-707 du 10/07/1964 : création des départements 75 (Paris), 78 (Yvelines), 91 (Essonne), 92 (Hauts-de-Seine), 93 (Seine-Saint-Denis), 94 (Val-de-Marne) et 95 (Val-d'Oise)
- loi du 29/12/1967 : rattachement de 6 communes de l'Ain et de 24 communes de l'Isère au département du Rhône
- décret du 10/10/1969 : les Basses-Pyrénées deviennent Pyrénées-Atlantiques
- décret du 13/04/1970 : les Basses-Alpes deviennent Alpes-de-Haute-Provence
- loi n° 75-356 du 15/05/1975 : création des départements 2A (Corse-du-Sud) et 2B (Haute-Corse)
- décret du 27/02/1990 : les Côtes-du-Nord deviennent Côtes-d'Armor

Liste des départements retenus pour le fichier départemental

Code Département	Code Département
01 Ain	48 Lozère
02 Aisne	49 Maine-et-Loire
03 Allier	50 Manche
04 Alpes-de-Haute-Provence	51 Marne
05 Hautes-Alpes	52 Haute-Marne
06 Alpes-Maritimes	53 Mayenne
07 Ardèche	54 Meurthe-et-Moselle
08 Ardennes	55 Meuse
09 Ariège	56 Morbihan
10 Aube	57 Moselle
11 Aude	58 Nièvre
12 Aveyron	59 Nord

13	Bouches-du-Rhône	60	Oise	
14	Calvados	61	Orne	
15	Cantal	62	Pas-de-Calais	
16	Charente	63	Puy-de-Dôme	
17	Charente-Maritime	64	Pyrénées-Atlantiques	
18	Cher	65	Hautes-Pyrénées	
19	Corrèze	66	Pyrénées-Orientales	
2A	Corse-du-Sud	Fusionnés en 20 Corse	67	Bas-Rhin
2B	Haute-Corse		68	Haut-Rhin
21	Côte-d'Or	69	Rhône	
22	Côtes-d'Armor	70	Haute-Saône	
23	Creuse	71	Saône-et-Loire	
24	Dordogne	72	Sarthe	
25	Doubs	73	Savoie	
26	Drôme	74	Haute-Savoie	
27	Eure	75	Paris (Seine avant 1968)	
28	Eure-et-Loir	76	Seine-Maritime	
29	Finistère	77	Seine-et-Marne	
30	Gard	78	Yvelines (Seine-et-Oise avant 1968)	
31	Haute-Garonne	79	Deux-Sèvres	
32	Gers	80	Somme	
33	Gironde	81	Tarn	
34	Hérault	82	Tarn-et-Garonne	
35	Ille-et-Vilaine	83	Var	
36	Indre	84	Vaucluse	
37	Indre-et-Loire	85	Vendée	
38	Isère	86	Vienne	
39	Jura	87	Haute-Vienne	
40	Landes	88	Vosges	
41	Loir-et-Cher	89	Yonne	
42	Loire	90	Territoire de Belfort	
43	Haute-Loire	91	Essonne (à partir de 1968)	
44	Loire-Atlantique	92	Hauts-de-Seine (à partir de 1968)	
45	Loiret	93	Seine-Saint-Denis (à partir de 1968)	
46	Lot	94	Val-de-Marne (à partir de 1968)	
47	Lot-et-Garonne	95	Val-d'Oise (à partir de 1968)	
		97	DOM	

Informations reprises et adaptées du cédérom d'origine de l'INSEE.