1. Qui a créé ce jeu de données ? À quelles fins ?

La base de données des prénoms a été créée en 1946 par l’INSEE, l'Institut national de la statistique et des études économiques (INSEE), un service central du Ministère chargé des Finances lui-même créé en 1946. Cette institution détient une indépendance vis-à-vis de tout organe gouvernemental depuis sa création. Actuellement, son directeur est Jean-Luc Tavernier. Chaque année, l’Insee collecte, produit, analyse et diffuse des informations sur l'économie et la société françaises. Son devoir principal est de rester neutre et de ne pas interpréter les données qu’elle collecte. Parmi ses nombreuses fonctions statistiques et régaliennes, elle a pour mission de recenser la population française et d’organiser ces données afin de les publier pour qu’elles soient exploitées. Parmi ses études nationales, l’Insee s’intéresse aussi aux prénoms attribués aux enfants nés chaque année.

C’est ainsi que tous les ans depuis 1946, l’INSEE publie un fichier des prénoms basé sur les bulletins d’état civil transmis par les officiers d’état civil de chaque commune. L’INSEE a fait un travail de récolte de données d’état civil en remontant jusqu’à l’année 1900, ce qui fait de cette base un outil d’une valeur historique et sociologique inestimable. Ces bulletins sont remplis par l’officier d’état civil sur déclaration des parent au moment de la déclaration de naissance d’un enfant, et sont conservés dans les mairies des communes. Le fichier se présente comme une liste de prénoms, à côté desquels est affiché toutes les années pour lesquels ces prénoms ont été donnés au moins trois fois dans la même année dans l’ordre chronologique. Les années pour lesquelles les données ne sont pas suffisantes ne sont pas inscrites : ainsi, à côté de certains prénoms, ne figurent qu’une seule ou quelques années seulement. Cette base de données concerne l'ensemble des personnes nées en France excepté Mayotte. Sont donc exclus les enfants français nés à l’étranger, et les enfants étrangers nés en France sont en revanche inclus.

Depuis 1946, la base de données des prénoms de l’INSEE se base sur le répertoire national d’identification des personnes physiques (RNIPP), dont l’INSEE est également responsable, et qui est, un *“instrument de vérification de l’état civil des personnes nées en France”*. Cet outil très util et très précis recense les noms et prénoms, lieux et date de naissance, sexe, date et lieu de décès, ainsi que les numéros des actes correspondant, et les éventuelles modifications d’état civil successives. Les données sont conservées même après la mort de la personne. Ainsi, dans son état actuel, le fichier des prénoms de l’INSEE est un “sous-RNIPP”, puisque lui ne prend en compte qu’une des données de cette base de données colossale, les prénoms, sans aucune autre précision que l’année au cours de laquelle ils ont été donnés. Le RNIPP a été extrêmement important pour le suivi des prénoms donnés en France, puisqu’il a permis de simplifier l’actualisation et le suivi du fichier des prénoms de l’INSEE. L’INSEE explique elle-même, sur son site internet, que les données recueillies avant 1946 sont parcellaires, en tout cas pas tout à fait exhaustives, et que les utilisateurs de la base de donnée pourront donc *“constater un écart entre le nombre annuel des naissances calculées par l’INSEE”*. En effet, pour les années avant 1946, l’INSEE a mené un travail de rédolte de données auprès des mairies et de leurs officiers d’état civil. Ce fichier, bien que très complet, n’a pas vocation à être exhaustif. L’INSEE admet qu’il peut comporter certaines erreurs ou oublis, surtout avant 1946.

Le fichier de l’INSEE est un fichier public, ouvert à tous, et qui a une certaine vocation de transparence des administrations publiques et des actes qu’elles prennent. En outre, l’INSEE remplit à ce titre ses fonctions d’institut de la statistique et d’information publique sur la question. Ainsi, les données peuvent être utilisées par n'importe qui, du simple citoyen cherchant à savoir combien de personnes portant le même prénom qu’elle sont nées la même année que lui, aux chercheurs et scientifiques l’utilisant pour analyser des phénomènes historiques ou sociologiques, en passant par les entités étatiques pour la prise de décision. Par exemple, il est certain que le fichier des prénoms a été un outil plus qu’essentiel pour la création de la loi de 1993 visant à donner une plus grande latitude aux parents dans le choix du prénom de leurs enfants. Nous allons voir, tout au long de cette étude, que le fichier des prénoms de l’INSEE peut être utilisé pour interpréter nombre de phénomènes, et qu’il est extrêmement intéressant de se pencher un tant soit peu sur ce fichier.