

## Fiche d'animation - DataStory: les prénoms

Cette fiche détaille une animation de médiation aux données sous la forme d'une DataStory.

Le principe d'une DataStory est de s'appuyer sur un thème et de le décliner sous la forme d'ateliers d'une durée variable (de 1h à 4h). Chacun de ces ateliers permet de traiter un aspect de la médiation aux données:

- identifier et trouver des données pertinentes,
- aborder les données avec une lecture critique, aborder le principe des métadonnées,
- savoir lire une datavisualisation, la décrypter et en identifier les limites
- rendre tangible les données sous la forme d'une visualisation "grandeur réelle"

Cette fiche présente un exemple de DataStory sur le thème des prénoms, mais la communauté Infolab est invitée à s'en saisir pour proposer d'autres thèmes.

### Les ressources

Afin de construire votre propre DataStory, nous vous conseillons de commencer par identifier des ressources pertinentes, qu'il s'agisse de données, de vidéos, d'articles de presse, de billets de blog, etc. L'objectif est de mieux cerner le sujet et de construire peu à peu votre kit de ressources.

Sur le thème des prénoms, nos ressources comprennent:

- des vidéos d'archives proposées par l'INA (visualisation gratuite, téléchargement payant mais à très faible coût)
  - "[La Saint Méthode](#)": un reportage de 1964 sur un prénom original, Méthode comprenant un micro-trottoir,
  - "[Attention aux prénoms farfelus](#)": un extrait d'un journal télévisé de 2003 à l'occasion des dix ans de la dernière loi modifiant les règles de l'état-civil
- une collection de datavisualisations que nous avons rassemblé en un point unique pour permettre un accès plus rapide: [checkthis.com/passionprenoms](http://checkthis.com/passionprenoms) - cette collection a été constituée en effectuant une recherche d'image sur Google avec les termes "prénoms", "visualisation", "baby names", ...
- des articles de la presse locale et régionale que nous avons collecté en bibliothèque. Le premier bébé de l'année est un "marronnier" classique de la première édition de l'année (entre le 2 et le 5 janvier). Ces articles font très souvent état du palmarès des prénoms des enfants nés l'année précédente.
- le [blog](#) de Baptiste Coulmont, sociologue de l'Université Paris 8, qui est aussi l'auteur de l'ouvrage de référence sur le sujet (Sociologie des prénoms, Éditions La Découverte) et notamment une [présentation synthétique](#) du sujet,
- quelques articles scientifiques repérés au fil de l'eau, notamment à l'aide de Google Scholar.
- certains ateliers déjà réalisés ont été documentés sur le blog donneesouvertes.info: "[Oups, on a fait un infolab](#)" et "[infolab #3, de la mode et des prénoms](#)"

Ce kit a été construit au fur et à mesure des animations, sur une période de 3 ans, et avec la complicité de Loïc Haÿ. Il est bien sûr possible de commencer à travailler un nouveau thème sans avoir constitué un kit aussi complet dès le démarrage. En revanche, il est important de penser à capitaliser toutes ces ressources pour en permettre le partage. Retrouvez ce document sur internet, notamment pour en suivre les hyperliens, à l'adresse: <http://infolabs.io/datastory-les-prenoms>

DataStory prénoms <http://infolabs.io/datastory-les-prenoms>

Par Simon Chignard et l'équipe Campagne Infolab, avec le soutien de Loïc Haÿ  
Licence Creative commons 3.0 Attribution France.

## Module 1 - “Découvrir le sujet et identifier des questions”

45 minutes

Tout public à partir de 12 ans<sup>1</sup>

Nombre de participants : 5 minimum / 25 maximum

Matériel: 1 ordinateur et vidéoprojecteur, 1 paper board,

Connexion Internet : Indispensable

Ce module propose une introduction au sujet “Prénoms” via des ressources multimédias (vidéos d’archives et articles de presse).

On commence par la diffusion des deux vidéos de l’INA (cf. ressources). Après chaque vidéo, on demande au groupe de les résumer et on synthétise sous la forme “Qu’est-ce que cette vidéo nous raconte sur les prénoms ?”.

La vidéo sur la Saint Méthode (1964) permet ainsi d’évoquer :

- les prénoms rares ou originaux,
- le prénom comme indicateur d’une classe sociale - verbatim du micro-trottoir: “Méthode ? c’est trop haut placé pour moi” [0’36],
- l’influence supposée des stars ou des héros populaires dans le choix des prénoms “je ne sais pas si les parents se sont inspirés des vedettes du moment”, officier de l’état-civil [1’20],
- ...

La vidéo du JT de France 2 (2003) met en garde contre les prénoms farfelus. Elle permet d’évoquer:

- le libre choix des parents depuis la loi de 1993 [début du reportage],
- l’influence supposée des marques “Loreal, Chanel, Ikea”...

En fonction du temps disponible, on peut aussi compléter par des articles de presse (cf. ressources).

L’ensemble permet d’identifier deux questions sur le choix des prénoms, questions que l’on pourra ensuite traiter avec les données [cf. module 3]:

- le prénom comme une mode: les parents sont-ils totalement influencés par les prénoms des stars, des noms de marque, ... ?
- la recherche de l’originalité à tout prix: tout le monde ne s’appelle pas Arthur ou Marie...

---

<sup>1</sup> Pour les plus jeunes, on recommande de s’inspirer de l’atelier “[la dataviz et les enfants](#)” proposé par [Léa Lacroix](#)

## Module 2 - “Rechercher des données et savoir les lire”

45 minutes

Tout public à partir de 12 ans

Nombre de participants: 5 minimum / 25 maximum

Matériel: au moins 1 ordinateur pour deux participants (idéalement ordinateur par personne), 1 vidéoprojecteur pour l'animateur.

Connexion Internet : Indispensable.

L'objectif du second module est d'aborder la recherche de données et la notion de méta-données. On pourra aussi aborder le droit d'accès aux documents administratifs.

En lien avec le parcours Data et l'interview de la donnée.

En pratique, on demande aux participants d'identifier les prénoms des enfants nés à Rennes, Paris ou Nantes en 2012 [toutes ces données existent en ligne, sur les portails open data des villes concernées] :

- [http://www.data.rennes-metropole.fr/les-donnees/catalogue/?tx\\_icsopendatastore\\_pi1\[keywords\]=&tx\\_icsopendatastore\\_pi1\[categories\]\[0\]=14&tx\\_icsopendatastore\\_pi1\[uid\]=169](http://www.data.rennes-metropole.fr/les-donnees/catalogue/?tx_icsopendatastore_pi1[keywords]=&tx_icsopendatastore_pi1[categories][0]=14&tx_icsopendatastore_pi1[uid]=169)
- <http://data.nantes.fr/donnees/detail/prenoms-enfants-nantes/>
- [http://opendata.paris.fr/explore/dataset/liste\\_des\\_prenoms\\_2004\\_a\\_2012/?tab=table](http://opendata.paris.fr/explore/dataset/liste_des_prenoms_2004_a_2012/?tab=table)

Une fois la ressource identifiée, on laisse 5 minutes aux participants pour en prendre connaissance, puis on revient vers eux.

- qui a produit ce fichier ?
- est-ce qu'il y a tous les enfants de la ville ou seulement une partie d'entre eux ?
- qu'est-ce que j'ai le droit de faire avec ces données ?
- quand le fichier a-t-il été publié ? qui l'a publié ?

Les questions portent exclusivement sur les méta-données car, dans 95% des cas, les participants auront directement consulté le fichier lui-même sans prêter attention à la fiche de méta-données qui l'accompagne. Le fichier concerné ne répond pas à ces questions.

La question “qu'ai-je le droit de faire avec ces données ?” est aussi l'occasion d'aborder la notion de licence de réutilisation:

- qu'est ce qu'une licence de réutilisation de données ?
- quelles sont les droits et les obligations des réutilisateurs ?
- comment “attribuer” une donnée à sa source ?

On pourra se reporter aux textes de la [licence ouverte](#) et de la licence [ODbL](#), cette dernière ayant le mérite d'être résumée sous la forme de pictogramme.

En fonction du temps dont l'on dispose, on peut aussi aborder la question du secret statistique.

[Ressource sur le sujet prénoms et secret statistique lors de la journée Open Data à la Cnil: [page 4](#)]

## Module 3 - “Les prénoms dans ma ville”

60 à 90 minutes

Tout public à partir de 12 ans

Nombre de participants: 5 minimum / 25 maximum

Matériel: au moins 1 ordinateur pour deux participants (idéalement ordinateur individuel), 1 vidéoprojecteur pour l'animateur, 1 fichier prénoms sur chaque ordinateur ou en version papier pour 2 participants..

Connexion Internet : indispensable.

Cette animation nécessite de disposer d'un fichier sous format numérique des prénoms donnés aux enfants d'une ville sur une ou plusieurs années. Ce fichier aura pu être identifié lors du module 2, ou être fourni sous format imprimé aux participants [recommandé si l'on dispose de moins de temps]

Chaque participant est d'abord invité à prendre connaissance du fichier pendant 5 minutes. Puis ensuite on invite tous les participants à faire leur rapport d'étonnement et à poser leurs questions.

Parmi les remarques et questions déjà entendues:

- “il n'y a pas de prénoms composés”
- “il y a tous les enfants, là ?”
- “le prénom Théo / Maélys s'écrit avec plusieurs orthographes: ça compte une fois ou plusieurs fois ?” ...

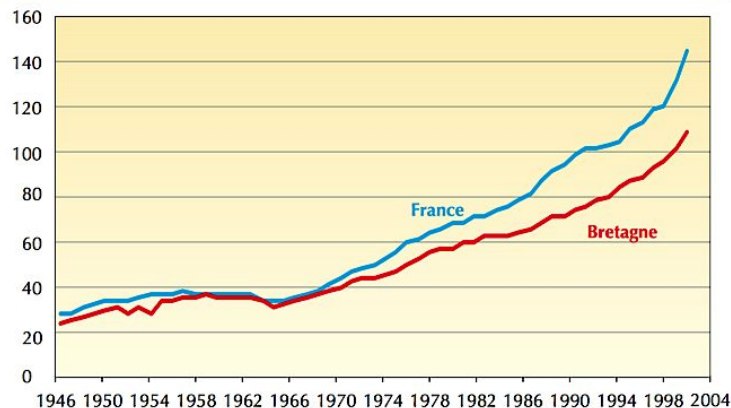
On peut ensuite reprendre les motifs identifiés dans le module 1 (influence de la mode, recherche de l'originalité) et essayer de les confirmer à partir des données.

Si l'on dispose d'une version numérique et d'un ordinateur équipé d'un logiciel tableur (Open Office, Excel) et que l'on connaît le nombre total de naissances, on peut notamment calculer :

- la part des naissances portant le prénom le plus fréquent (le premier),
- la part des naissances portant l'un des 10 prénoms les plus fréquents,
- le nombre de prénoms nécessaires pour nommer la moitié d'une classe d'âge.

Ce dernier indicateur a fortement évolué dans le temps: il montre bien la grande diversité des prénoms et la recherche de l'originalité qui prévaut. Cf. graphique issu de la [revue Octant](#) (Insee Bretagne)

### Évolution du nombre de prénoms nécessaires pour nommer la moitié d'une classe d'âge



Lecture : de 1946 à 1970, moins de 40 prénoms suffisaient à nommer la moitié des enfants nés en Bretagne ; en 2002, il en faut 109.  
Source : Insee

DataStory prénoms <http://infolabs.io/datastory-les-prenoms>

Par Simon Chignard et l'équipe Campagne Infolab, avec le soutien de Loïc Haÿ  
Licence Creative commons 3.0 Attribution France.

## Module 4 - “Lire une datavisualisation et en comprendre les limites”

45 minutes

Tout public à partir de 12 ans

Nombre de participants : 5 minimum / 25 maximum

Matériel: 1 ordinateur avec vidéoprojecteur pour l’animateur,

Connexion Internet : indispensable.

A partir de la [collection](#) de datavisualisations sur les prénoms (cf. ressources).

Ce module vise à développer une lecture critique des visualisations de données. Pour chaque exemple présenté, on demande aux participants:

- que représente cette image ? qu’est-ce qui vous étonne ?
- quel est son titre, sa légende ?
- à quoi ressemblent les données sources ? d’où viennent-elles ?
- quels choix a fait celui a crée cette image ?
- quels sont les codes graphiques retenus ?
- quelles sont les limites de cette représentation ? comment l’améliorer ?

Exemple sur la première image: “Europe’s most popular given female names”



Cette [carte](#) d’Europe représente le prénom féminin le plus populaire dans chaque pays. On ne connaît pas la date (données annuelles ? sur plusieurs années ? sur un seul mois ?), ni la source.

On imagine que le fichier source comprend deux colonnes: le nom du pays et le prénom le plus populaire. En première lecture on voit que plusieurs pays ont des prénoms en commun, par exemple Emma en France, Belgique et Luxembourg. On note aussi le très grand nombre de prénoms finissant par la lettre A (Elena, Lucia, Maria, Olivia, ...) et les exceptions (Nino, Gudrun).

Le code couleur ne semble pas obéir à une logique particulière, hormis de marquer les frontières entre pays. La limite principale de ce graphique est de montrer uniquement le prénom le plus populaire. Or ce prénom ne représente qu’une toute petite part des naissances (1 à 3% en moyenne). C’est un peu comme représenter un thème comme “l’impôt et les Français” en ne regardant que ceux qui paient à l’impôt de solidarité sur la fortune !

DataStory prénoms <http://infolabs.io/datastory-les-prenoms>

Par Simon Chignard et l’équipe Campagne Infolab, avec le soutien de Loïc Haÿ  
Licence Creative commons 3.0 Attribution France.

## Module 5 - “Faire une dataviz grandeur nature”

4 heures

Tout public

Nombre de participants adapté : 5 minimum / 10 maximum

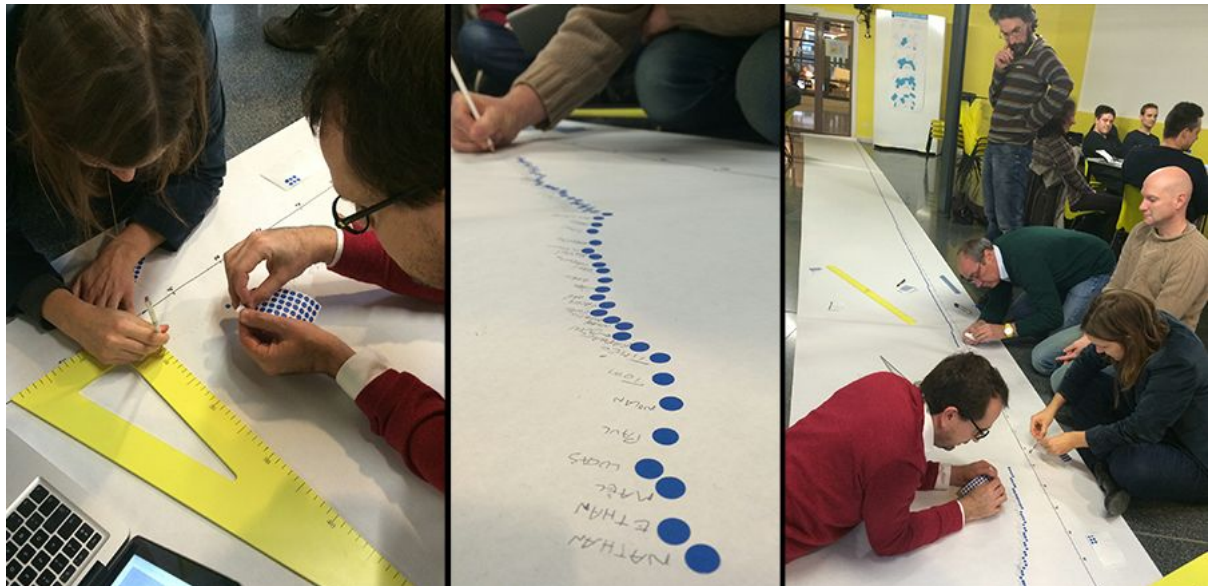
Matériel: 1 ordinateur, fichier au format numérique des naissances sur une ville et une année, 1 rouleau de 25 m de papier (nappe), autant de gommettes que de prénoms différents ou de naissances, des feutres, du matériel pour tracer des grandes lignes, une grande équerre graduée et une grande règle (disponible dans les magasins de fourniture pour écoles ou de loisirs manuels).

Connexion Internet : facultative

La distribution des prénoms se caractérise par un phénomène de type [longue traîne](#): un très grand nombre de prénoms ne sont donnés qu’une à 3 fois dans une même ville chaque année [plus exactement, il s’agit d’une distribution obéissant à la [loi de Zipf](#)]. La recherche de l’originalité est devenue la norme. Comment rendre tangible ce phénomène ?

Dans le cadre de la journée Infolab à Rennes en décembre 2013, nous avons réalisé une fresque grandeur nature, d’une quinzaine de mètres de long. Cette fresque a été conçue pour représenter en abscisse le nombre d’enfants qui porte un prénom et en ordonnée chaque prénom classée par fréquence décroissante. On rajoute ensuite à la main le prénom à côté de la gommette.

Exemple: la première gommette correspond au prénom Nathan, qui a été donné 35 fois au cours de l’année considérée.



En amont il faut préparer le fichier et évaluer notamment la dimension attendue de la fresque. On répartit ensuite les rôles, qui seront amenés à changer au cours de l’atelier: certains sont chargés de tracer l’abscisse et l’ordonnée (deux traits), un autre lit la donnée sur l’ordinateur portable, un dernier pose la gommette et inscrit le prénom.

Cet atelier permet de donner une dimension sensible à la donnée. Et elle l’est d’autant plus que la réalisation d’une telle fresque se révèle vite fastidieuse, d’où l’importance de varier les rôles. Il est important de prévoir un espace pour exposer la fresque à l’issue de l’atelier. Lors de l’infolab rennais, elle constituait l’un des éléments mis en scène pour le public. Chacun essaye alors de rechercher son prénom - ou celui d’un proche - parmi toutes les gommettes.

DataStory prénoms <http://infolabs.io/datastory-les-prenoms>

Par Simon Chignard et l’équipe Campagne Infolab, avec le soutien de Loïc Haÿ  
Licence Creative commons 3.0 Attribution France.