

# **Master politiques sociales**

**Culture générale de la dataviz**

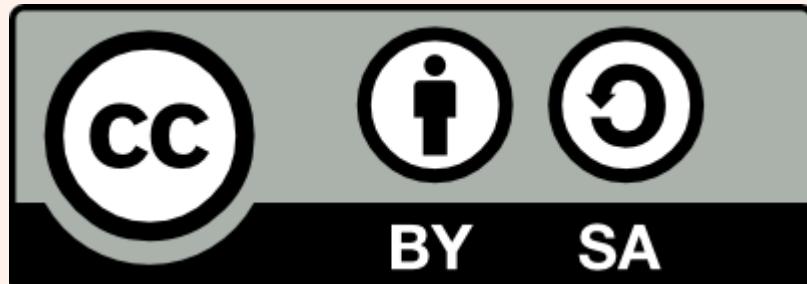
**Anne-Laure Donzel – Dataactivist**

**21/02/2024**

Ces slides en ligne : <http://dataactivist.coop/UAPV>

Sources : <https://github.com/dataactivist/UAPV>

Les productions de Dataactivist sont librement réutilisables selon les termes de la licence [Creative Commons 4.0 BY-SA](#).



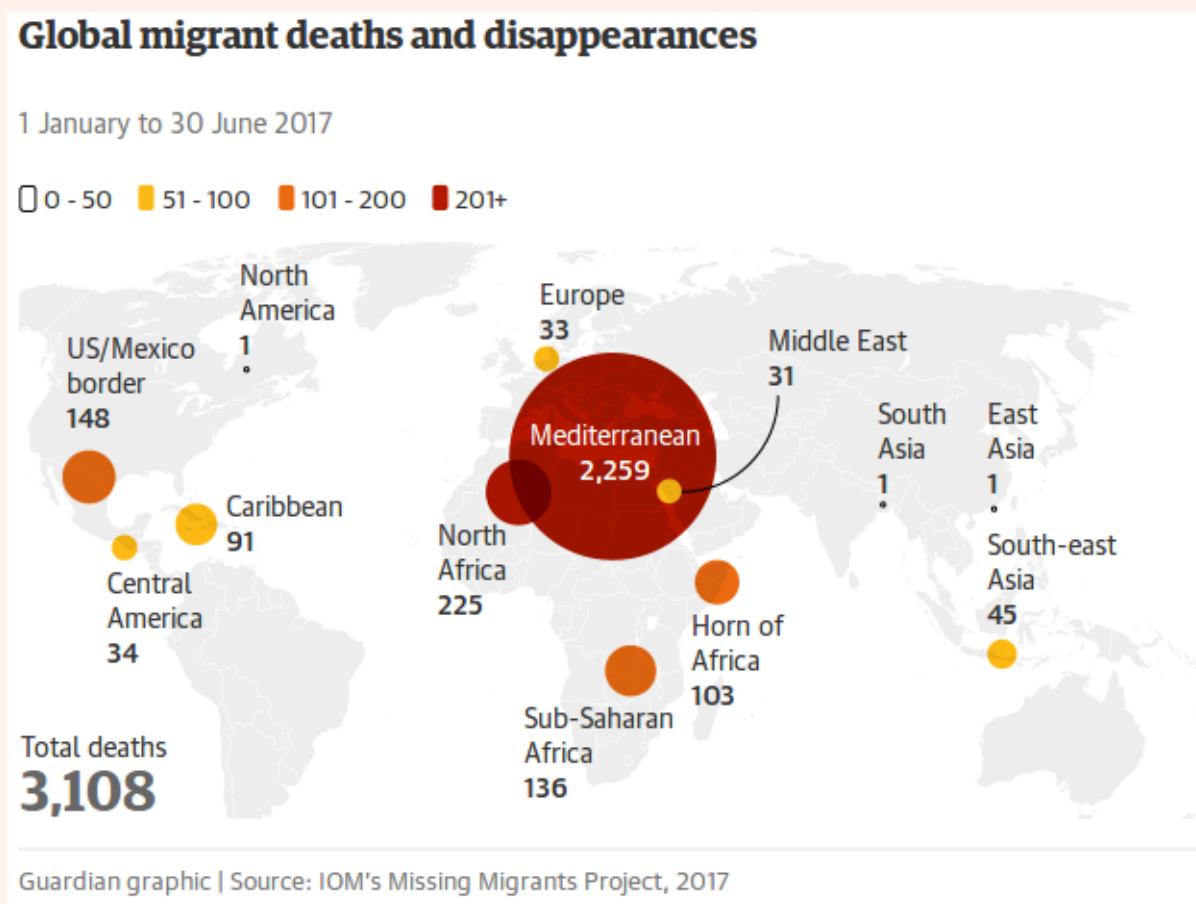
# **Comprendre la démarche de visualisation de données**

# Commençons par un exercice de visualisation

Selon l'Organisation internationale des migrations, 3108 migrant·e·s ont trouvé la mort ou sont disparu·e·s durant le premier semestre 2017.

- 2259 en Méditerranée
- 225 en Afrique du Nord
- 148 à la frontière américano-mexicaine
- 136 en Afrique subsaharienne
- 103 dans la Corne de l'Afrique
- 91 dans les Caraïbes
- 45 en Asie du Sud-Est
- 34 en Amérique centrale
- 33 en Europe
- 31 au Moyen-Orient
- un·e en Amérique du Nord
- un·e en Asie du Sud
- un·e en Asie de l'Est.

# Une autre représentation



*The Guardian*, "Migrant sea route to Italy is world's most lethal", 11/09/2017

# Que s'est-il passé ?

- Notre cerveau a tenté de se **représenter** les données qui nous étaient fournies. Il a utilisé ses capacités *d'imagination*.
- La carte propose une **représentation** visuelle des mêmes données.
- Cette représentation s'avère cependant plus fidèle, plus lisible et plus éloquente. Elle fait le travail qu'a tenté d'accomplir notre cerveau mais en y ajoutant des précisions, une clarté et un message **hors de sa capacité de traitement**.
- La datavisualisation agit donc comme un **outil cognitif** pour nous aider à comprendre les données.
- C'est ce qu'*Alberto Cairo* nomme la **valeur fonctionnelle** de la datavisualisation.

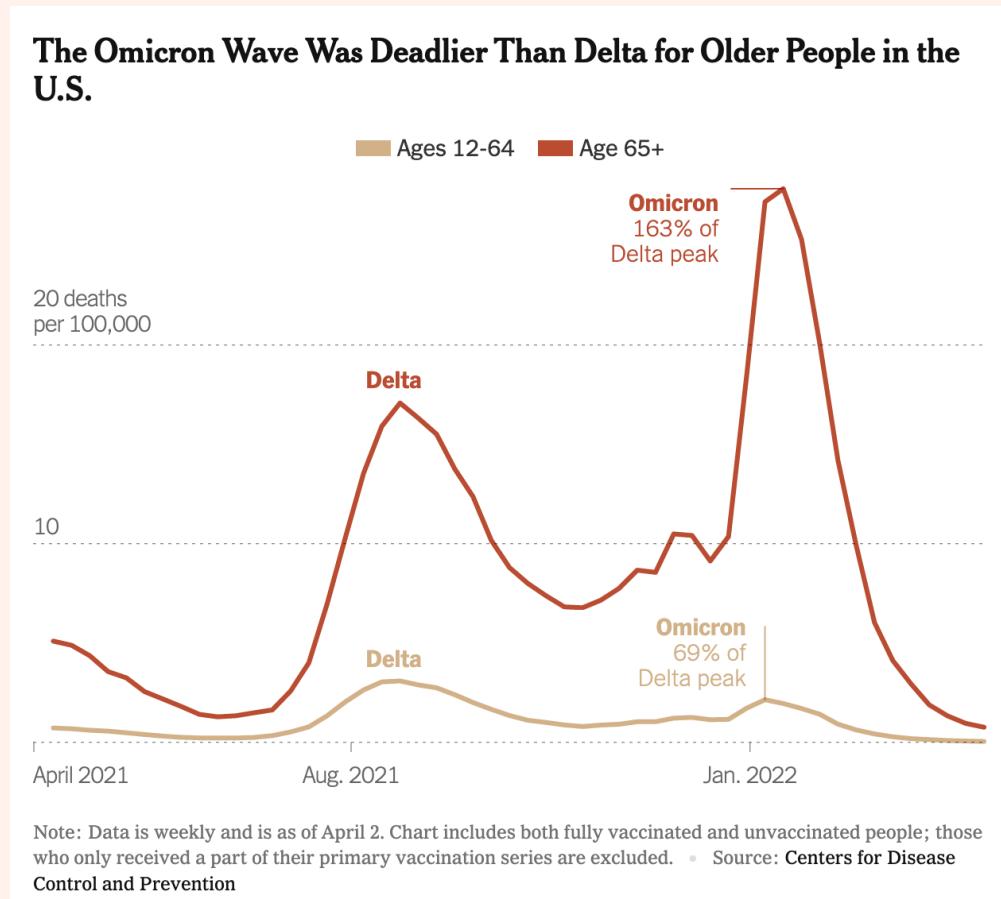
Alberto Cairo, *The functional Art*, 2012 ; *The truthful art*, 2016. À retrouver [sur son site web](#).

# Démarche d'encodage / décodage

- La visualisation de la donnée consiste en un **encodage**, c'est-à-dire une "conversion" d'un format à un autre.
- En l'occurrence, le passage d'une forme brute non interprétée (les *data*), à une forme raffinée interprétée (la *dataviz*).
- Comme dans le journalisme, la mise en forme implique également un **angle**, un *message* que doit transmettre l'image.
- Nous avons donc trois ingrédients indispensables à la démarche :
  1. des données ;
  2. une mise en forme ;
  3. un message.

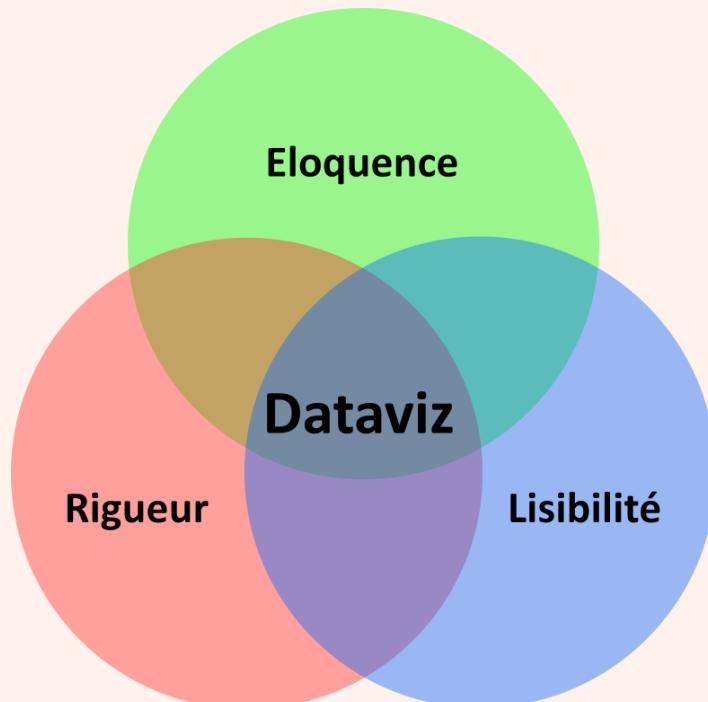
# Un exemple simple

Le *New York Times* a publié un article sur l'impact du Covid-19 sur les plus de 65 ans.



# **Qu'est-ce qu'une bonne dataviz ?**

**Trois qualités indispensables**



# Premier objectif : la rigueur

# Comprendre la nature et le sens des données

- La dataviz étant un encodage, elle substitue aux données des équivalences visuelles.
- L'encodage doit permettre :
  1. de comprendre **la nature** des données ;
  2. d'apprécier **le rapport** qu'elles entretiennent entre elles ;
  3. de saisir les points saillants et **phénomènes** clefs ;
  4. le tout sans déperdition de sens.
- Chacun de ces aspects repose sur des caractéristiques visuelles liées à des représentations.

# Expliciter la nature de la donnée



(source)

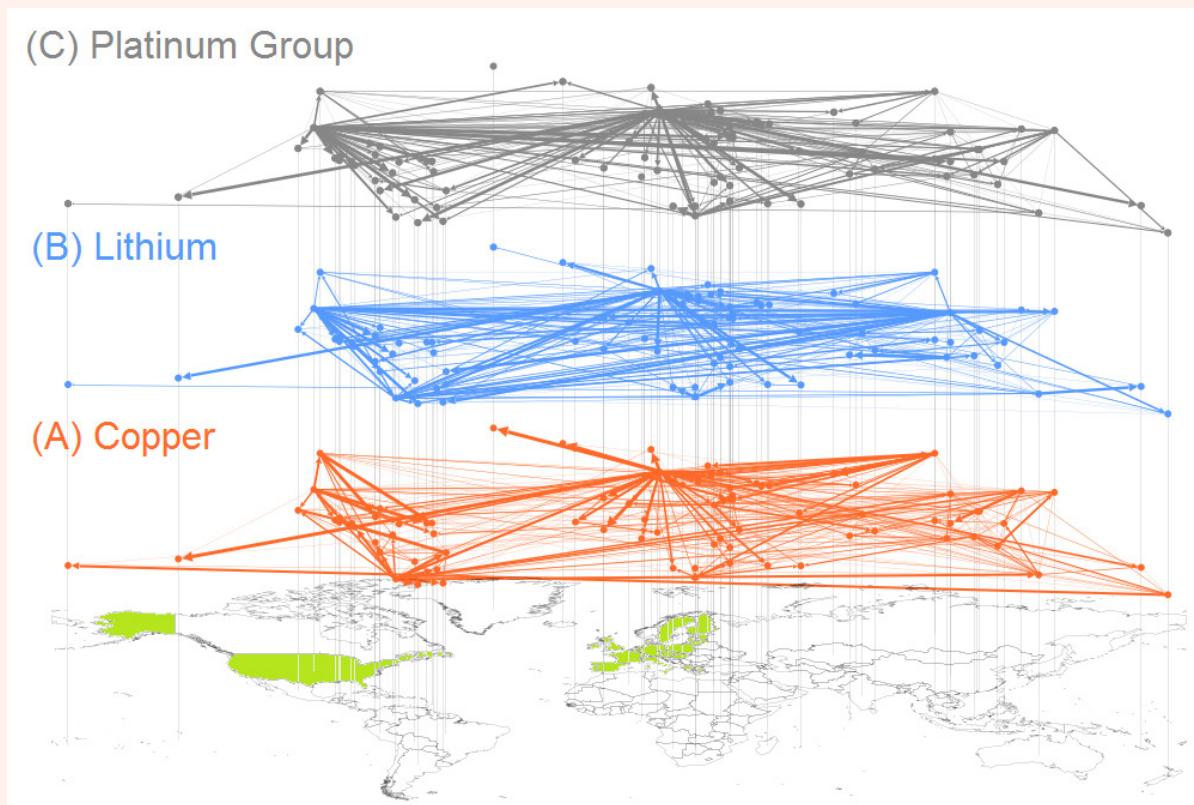
L'encodage doit permettre de saisir immédiatement le sens de chaque élément :

- où sont les points de données ?
- que représentent-ils individuellement / collectivement ?

La dataviz explicite ici le phénomène lui-même. Il s'agit du point de départ de la démarche journalistique. Les causes, elles, sont explorées et détaillées dans le texte.

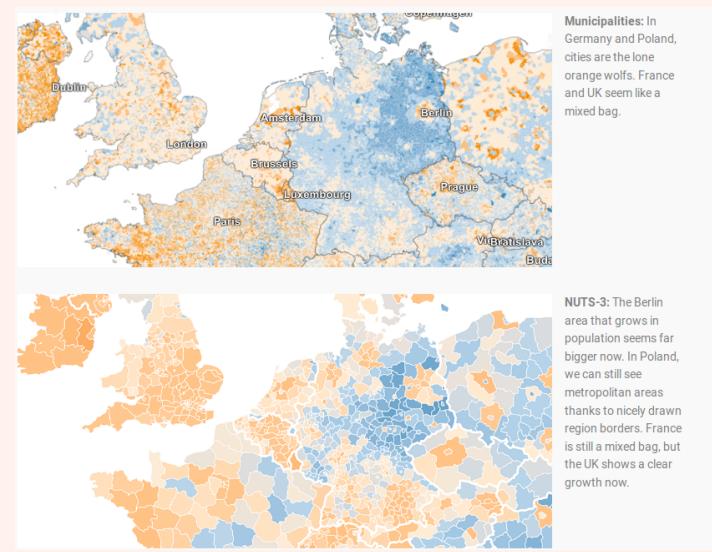
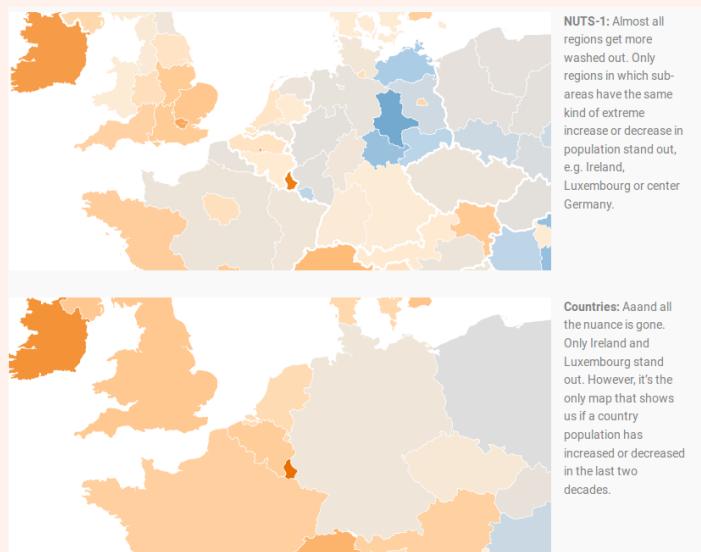
# Choisir la représentation adaptée au phénomène

Un mauvais choix graphique peut fausser la représentation de la donnée en ne donnant pas à constater le phénomène qu'elles permettent de décrire.



# Choisir le bon niveau de granularité

Le niveau de précision minimale d'une donnée (aussi appelé **granularité**) peut modifier de façon radicale la lecture d'une phénomène. Dans le cas d'une carte, le choix de l'échelle peut mener à la généralisation de phénomènes extrêmement circonscrits (ou inversement) comme le montre cet [exemple sur le blog de Datawrapper](#)

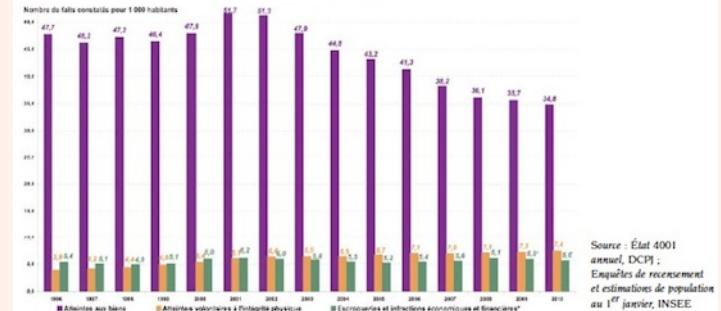


# Des graphs ni faits ni à faire

Le choix de l'échelle est aussi celui de son étendue. Il peut être tentant pour amplifier un phénomène d'en augmenter l'amplitude artificiellement en réduisant l'échelle. L'exemple le plus courant est le "["data-does-not-start-at-zero"](#)". Si ce choix peut se justifier, il est bien souvent utilisé comme méthode de manipulation.

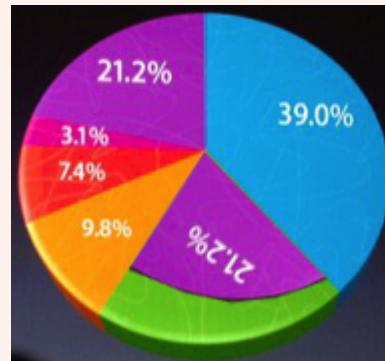
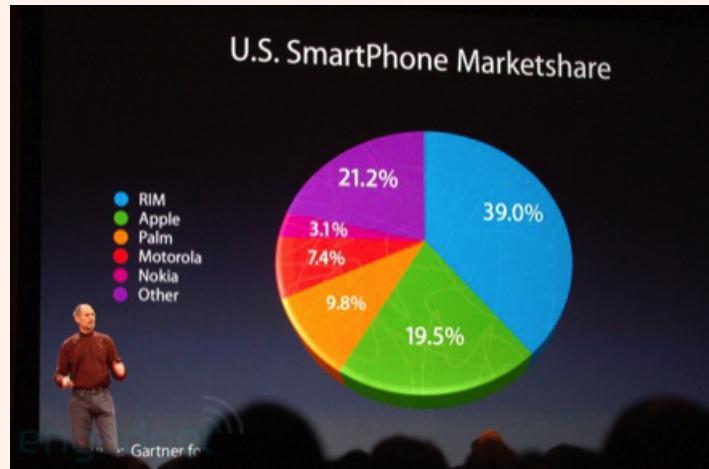


Graphique 2. Évolutions entre 1996 et 2010 du nombre de faits constatés d'atteintes aux biens, d'atteintes volontaires à l'intégrité physique et d'escroqueries et infractions économiques et financières pour 1 000 habitants en France métropolitaine, selon les enquêtes de recensement et les estimations de populations au 1<sup>er</sup> janvier de l'INSEE.



Jean-Marc Manach, "Plus la délinquance baisse, plus la violence augmente", *Owni*, Janvier 2011

# Des graphs ni faits ni à faire



Vandy Berten, "Comment mentir avec un graphique", Smals Research, Juillet 2015

# **Deuxième objectif : la lisibilité**

# 1<sup>er</sup> risque : "l'overplotting"

En statistique, le terme "overplotting" se réfère à l'effet produit par l'intégration d'un trop grand nombre de points de données dans un graph, le rendant illisible.

*Note : c'est un vrai terme de stat, vous pouvez l'utiliser pour vous la jouer.*

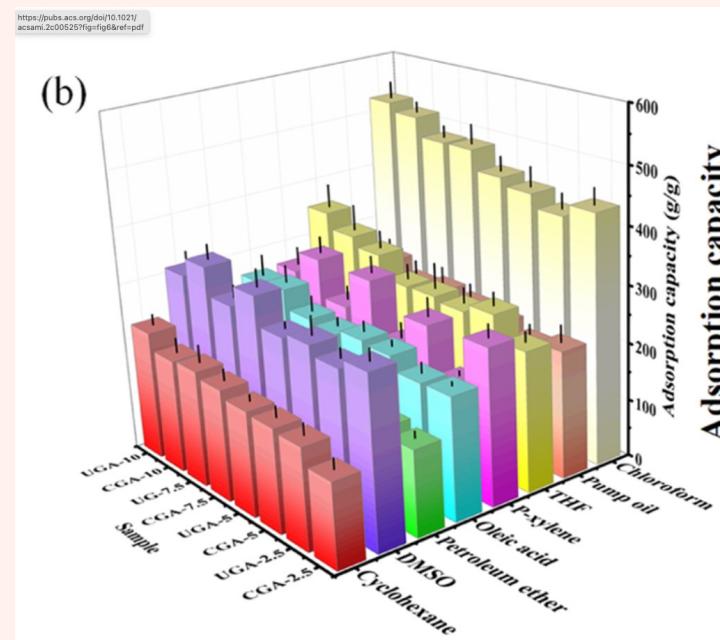
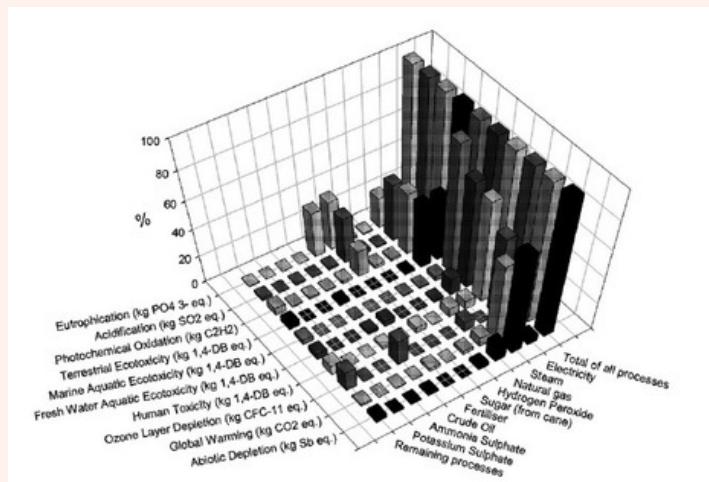
Exemple : une heat-map qui ne sert à rien



## 2<sup>e</sup> risque : "l'overcomplicated"

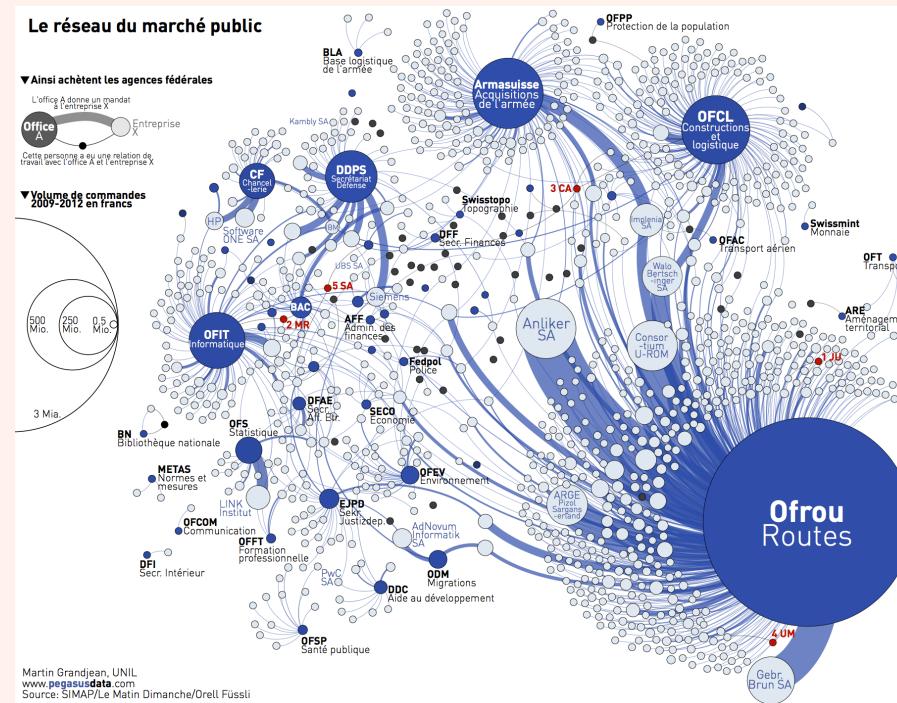
L'autre risque courant est de multiplier les dimensions et axes de lecture jusqu'à rendre le graphique incompréhensible.

*Note : ce terme n'a rien d'officiel, ne l'utilisez pas pour vous la jouer.*



# 3<sup>e</sup> risque : l'excès d'esthétisme

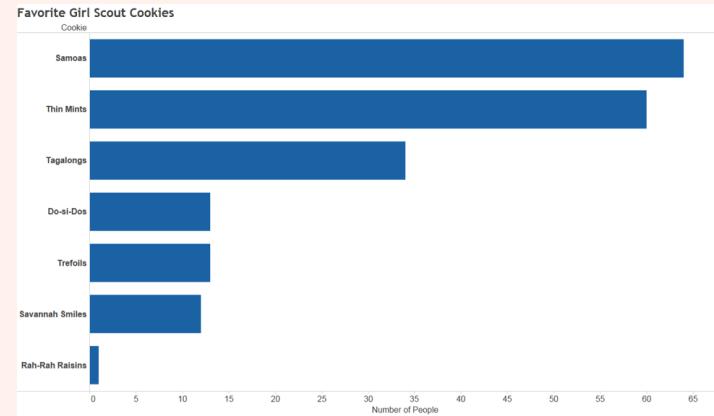
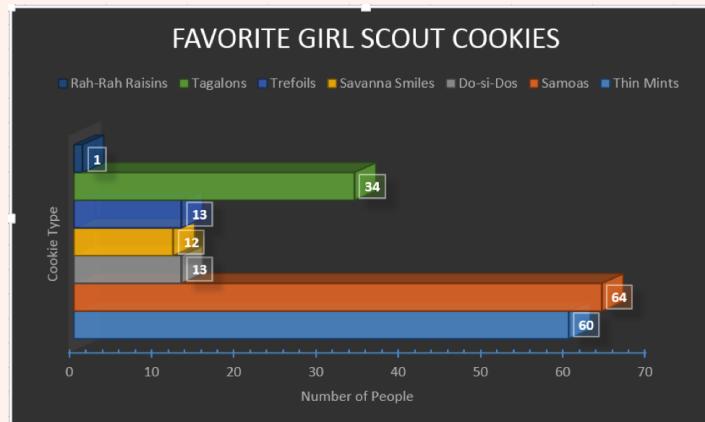
La tentation de mêler précision et esthétique peut produire de très beaux graphs... trop compliqués pour être compris du premier coup.



Martin Grandjean, Pegasus Data

# Laissez parler les données

Les fioritures peuvent constituer des distractions : quand les données sont claires, autant leur laisser le champ libre.



Nazirah Jetha, "5 data-viz tips to let your data speak for itself", Mai 2016

# Sans message, l'image devient vaine

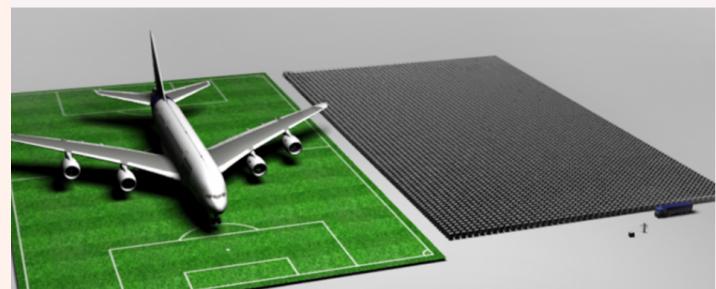
Il peut arriver que la dataviz n'ait d'intention que décorative ou spectaculaire. Précise et lisible, elle devient un panneau vide de sens et de propos qui n'explique rien du monde faute de choisir comment parler de son sujet.

Ex. : une "visualisation concrète de la dette française" qui n'a aucun sens

**Dette publique: 2.000 milliards d'euros, ça représente deux Stade de France en billets de 100 euros**

Fred Hasselot et Jean-Marie Pottier — 30 septembre 2014 à 13h50 — mis à jour le 30 septembre 2014 à 13h54

Une visualisation concrète de la dette de la France.



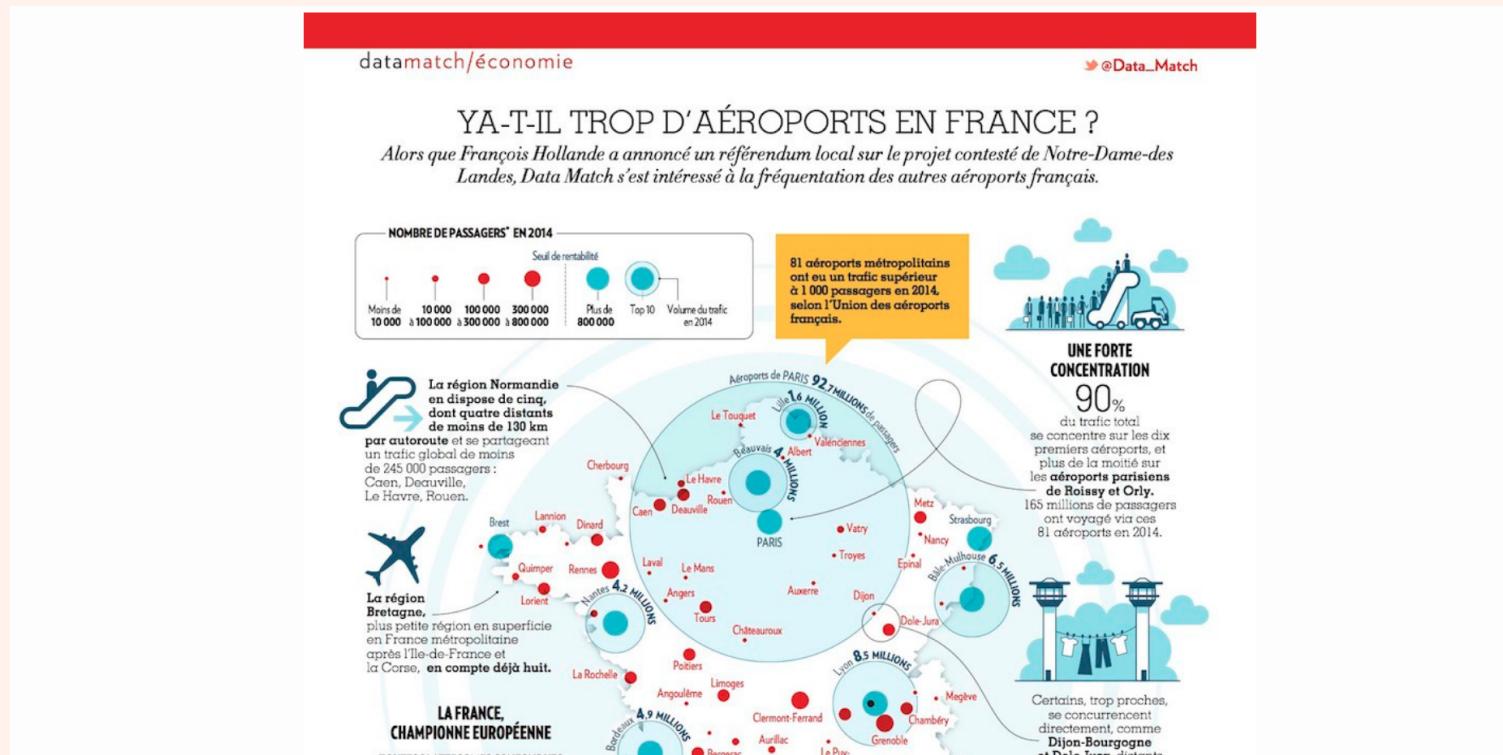
# **Troisième objectif :**

# **l'éloquence**

# Donner des points de repère

La façon la plus simple de mettre en avant un message reste encore de l'expliciter : souligner des points saillants ou relever les infos clefs permet de guider dans la lecture.

Ex. : rubrique DataMatch de Paris Match

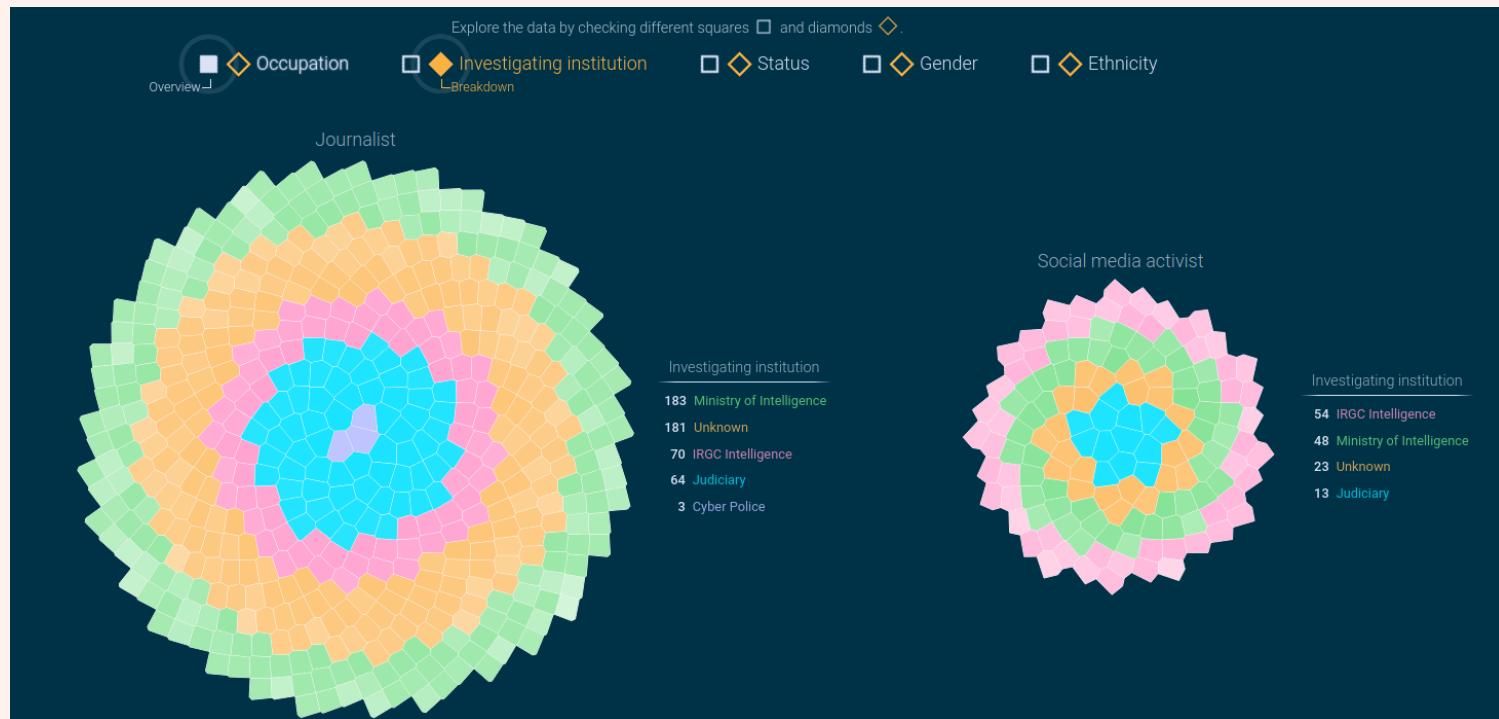


# **La symbolique comme message**

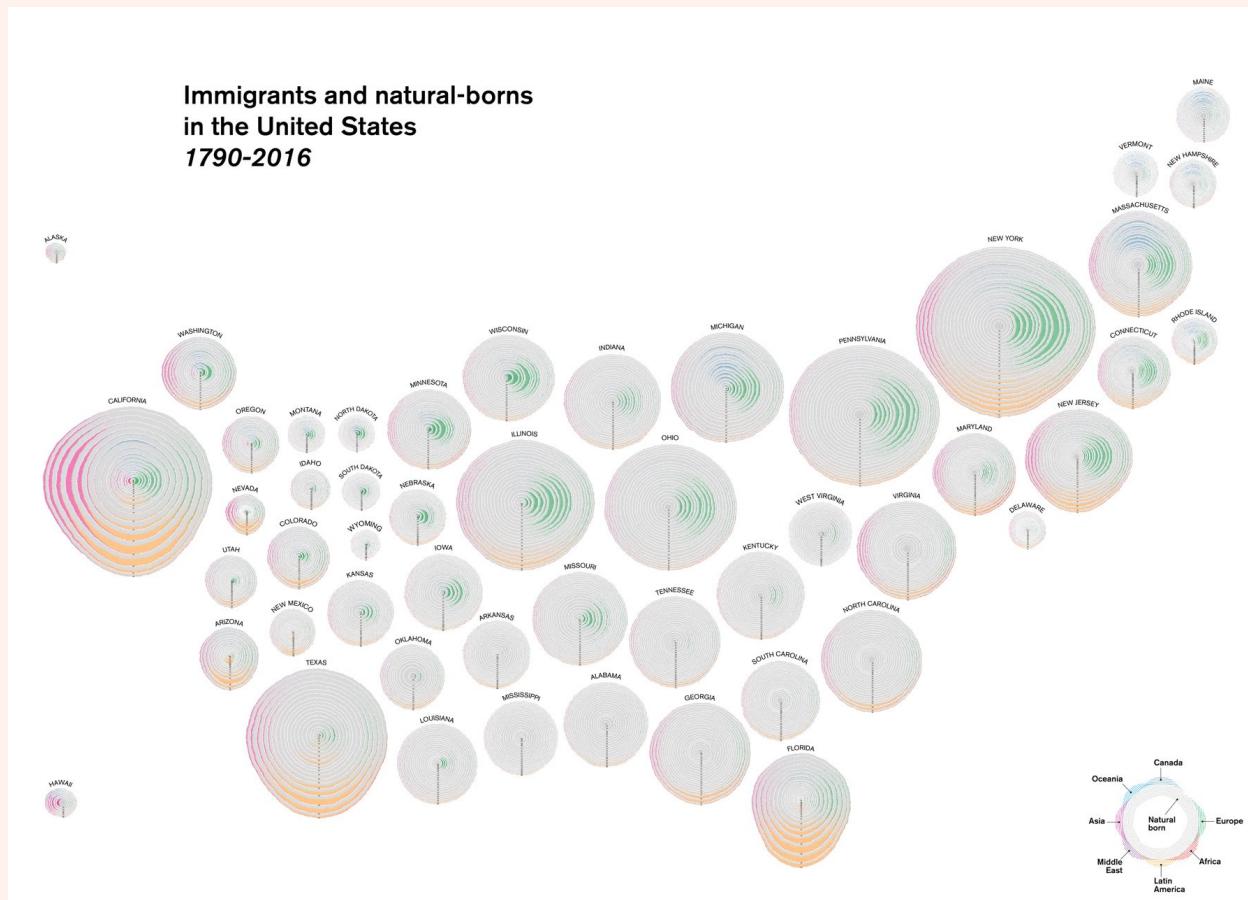
Le choix d'une représentation figurative allégorique plutôt que littérale peut constituer en soi l'angle d'une dataviz. Claire et bien choisie, la référence agit alors comme un sous-texte à l'image.

*Cf. exemples ci-après.*

Pour illustrer la censure en Iran, l'équipe de *Journalism is not a crime* a adopté l'esthétique des mosaïques d'Ispahan, faisant écho à l'imaginaire visuel du pays et à sa grandeur. Contrepied de la politique de censure menée par le régime ainsi dénoncé.



En proposant de visualiser l'immigration comme des anneaux de croissance plutôt que comme des flux, le chercheur en design d'information Pedro Cruz modifie la lecture du phénomène, basculant du registre visuel de la pression à une référence à la consolidation.



A sunburst chart of the US states, showing immigration and natural

# Permettre l'exploration

Des visualisations interactives peuvent inviter à l'exploration.

Ex. : *El País* a souhaité aborder la question du logement à Madrid en prenant le point de vue des habitant·es en recherche de logement.



# Faire appel à d'autres sens

Le message peut aussi passer par d'autres sens que la vue : l'ouïe, le toucher...

Ex. : le designer d'information Duncan Geere **a mis en son** une année de Covid-19 à Londres (une seconde = un jour) avec une ligne de basse pour les déplacements, deux notes séparées par un octave pour les cas et les décès, et des "pings" qui représentent des événements soudains (confinement, variant...).

