

LE MONDE DES DONNÉES



SFDS

LE NOMBRE ET LA CITÉ

*La statistique éclaire-t-elle
les questions de société ?*

Pont-Neuf

edp sciences

Le nombre et la cité

*La statistique éclaire-t-elle
les questions de société ?*

PONT-NEUF

Préface de Jean-Marie Delarue



17, avenue du Hoggar – P.A. de Courtabœuf
BP 112, 91944 Les Ulis Cedex A

Publications de la SFdS

Collection **Le Monde des données** (EDP Sciences)

Droesbeke J.-J. et Vermandele C. (2018), *Histoire(s) de(s) données numériques*.

Le Gléau J.-P. (2019), *Le secret statistique*.

Collection des **Journées d'étude en Statistique** (Éditions Technip)

Droesbeke J.-J., Maumy-Bertrand M., Saporta G. et Thomas-Agnan C. Éd. (2014), *Approches statistiques du risque*.

Droesbeke J.-J., Saporta G. et Thomas-Agnan C. Éd. (2015), *Méthodes robustes en statistique*.

Bertrand F., Droesbeke J.-J., Saporta G. et Thomas-Agnan C. Éd. (2017), *Model choice and model aggregation*.

Maumy-Bertrand M., Saporta G. et Thomas-Agnan C. Éd. (2018), *Apprentissage statistique et données massives*.

Collection **La statistique autrement** (Éditions Technip)

Droesbeke J.-J. et Vermandele C. (2016), *Les nombres au quotidien. Leur histoire, leurs usages*.

Ardilly P. et Lavallée P. (2017), *Les sondages pas à pas*.

Collection **Pratique de la statistique** (Presses Universitaires de Rennes)

Husson F. Éd. (2018), *R pour la statistique et la science des données*.

Bécue-Bertaut M. (2018), *Analyse textuelle avec R*.

Genuer R. et Poggi J.-M. (2019), *Forêts aléatoires avec R*.

Composition et mise en page : Patrick Leleux PAO

Couverture : Conception graphique de B. Defretin, Lisieux

Imprimé en France

ISBN (papier) : 978-2-7598-2374-1

ISBN (ebook) : 978-2-7598-2384-0

Tous droits de traduction, d'adaptation et de reproduction par tous procédés, réservés pour tous pays. La loi du 11 mars 1957 n'autorisant, aux termes des alinéas 2 et 3 de l'article 41, d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective », et d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation intégrale, ou partielle, faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (alinéa 1^{er} de l'article 40). Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles 425 et suivants du code pénal.

© EDP Sciences, 2019

Ce livre porte témoignage d'une série de rencontres, les Cafés de la statistique, dont l'ambition depuis quatorze ans est d'intéresser le citoyen à l'éclairage que la statistique donne aux questions de société.

Certains parmi les organisateurs de ces rencontres l'ont rédigé en commun : Brigitte Belloc, Alain Godinot, René Padieu, Sandra Roger, Jean-François Royer, Marion Selz et Olivier Vasseur. Ils ont choisi « Pont-Neuf » comme pseudonyme, parce que la plupart des Cafés se sont tenus dans une brasserie proche du Pont-Neuf à Paris.

Les Cafés de la statistique sont une activité du groupe « Statistique et enjeux publics » de la Société française de statistique.

À la mémoire de Dominique Allain (1956 – 2017), statisticienne, qui, dans sa pratique professionnelle comme au sein de la Société française de statistique a œuvré pour que cette discipline aide à éclairer les questions de société.

SOMMAIRE

<i>Préface : Nombre et société.</i>	11
<i>Avant-propos</i>	23

Partie 1

La statistique est intimement liée à la vie de la cité

1. La statistique aide à la compréhension du monde.	33
1.1 N'est-il de richesses que d'hommes ?	35
1.2 Nourrir 10 milliards d'humains	38
1.3 Matières premières et énergies : un désordre planétaire	42
1.4 Une consommation qui ignore les crises	45
2. La statistique repose sur des conventions et définitions partagées	51
2.1 « Entreprise » : qu'est-ce à dire ?	53
2.2 Territoires à la découpe	56
2.3 Les multiples facettes de la pauvreté	59
2.4 Les classes moyennes en quête de définition	63
3. La statistique permet de rationaliser des politiques publiques	67
3.1 Vieillir, oui, mais dans quel état et à quel coût ?	69
3.2 Faut-il construire des logements ?	73
3.3 Attachez votre ceinture !	76
3.4 Les déferlantes du tourisme	79

4. La statistique permet d'évaluer des politiques publiques	85
4.1 Recherche et développement : pour la croissance ?	87
4.2 Évaluation et recherche du consensus	90
4.3 Prévoir les bouchons... pour les empêcher !	94
5. La statistique est un des fondements de la démocratie	99
5.1 Demain, quelle retraite ?	101
5.2 La dette publique, une valeur montante	104
5.3 Capter l'opinion publique ou la forger ?	107
5.4 Manifestez, manifestez, il en restera toujours quelques chiffres ! .	111
6. La statistique n'a d'utilité que si elle est indépendante	115

Partie 2

La statistique doit répondre aux questions sur son utilisation

7. La statistique peut-elle traiter tous les sujets ?	125
7.1 Le casse-tête financier local	127
7.2 Controverse sur les statistiques ethniques	130
7.3 Connait-on l'immigration ?	133
7.4 Le jeune enfant : sujet peu observé	137
7.5 Suicide : drame individuel, énigme collective	140
8. La statistique rend-elle compte à temps des transformations sociales ?	145
8.1 La santé à tout prix ?	146
8.2 La nouvelle famille	150
8.3 Le travail, si familier, si méconnu	153
8.4 Le « halo » du bénévolat	158
8.5 Délocalisations : où est la France ?	161
9. La statistique peut-elle surmonter les différences culturelles, linguistiques, juridiques ?	165
9.1 Comparaisons européennes : la statistique serait-elle fédératrice ?	167
9.2 Payons-nous trop d'impôts ?	170
9.3 Réussite scolaire : évaluer les élèves, ou le système ?	174
9.4 Le progrès des sociétés est-il mesurable ?	177

10. Les produits de la statistique sont-ils à la disposition des citoyens ?	181
10.1 Vraies mesures et faux-semblants ?	183
10.2 Les journalistes à l'article du chiffre	186

Partie 3

La statistique doit gagner la confiance par des pratiques irréprochables

11. La statistique emploie-t-elle toujours des méthodes rigoureuses ?	195
11.1 Un bon chiffre, ça peut toujours servir.	197
11.2 Statistique et vérification scientifique	200
12. La statistique est-elle un danger pour les personnes ?	205
12.1 Vous n'avez pas le profil !	207
12.2 Données massives : danger massif ?	210
13. Peut-on empêcher l'instrumentalisation de la statistique ?	215
13.1 Réchauffement climatique : quand l'incertitude nourrit les certitudes	217
13.2 Délinquance : elle augmente ou elle diminue ?	219
13.3 À qui revient la valeur ajoutée ?	223
13.4 Chômage : de la cacophonie à la polyphonie ?	226
14. Quelle déontologie la statistique appelle-t-elle chez ceux qui la font ?	231
14.1 Dites-nous ce que vous pensez	232
14.2 L'éthique du démographe	235
14.3 Des données en masse, un rêve de statisticien	238
<i>Les Cafés de la statistique</i>	243
Liste des Cafés de la statistique depuis l'origine.	245
<i>Dictionnaire des sigles et acronymes</i>	251
<i>Index des sujets</i>	255

PRÉFACE : NOMBRE ET SOCIÉTÉ

Les riches réflexions que les « Cafés de la statistique » ont conduites, et dont une part trouve très heureusement son expression dans le livre qu'on va lire ici, ont un très large champ d'investigation. D'une manière ou d'une autre, beaucoup conduisent à une interrogation sur les rapports entre le nombre et la société qui peut aujourd'hui prendre la forme suivante : est-ce la fin de l'âge d'or de la statistique ? On esquissera en quelques mots une réponse, au moins dans le cadre français (l'auteur de ces lignes n'ayant guère de compétences pour élargir au-delà le champ de vision).

La naissance des outils statistiques et leur usage par des esprits curieux n'ont pas coïncidé, et de loin, avec une large diffusion des données et leur utilisation dans le débat social. De William Petty¹ à René Carmille², leur utilisation a eu des buts précis, destinés souvent à la gestion publique (comptage des hommes pour les mobiliser, recensement des biens pour les imposer) ou à des corporations

1. William Petty (1623-1687) est regardé comme l'un des pères de l'utilisation des statistiques en économie et en démographie.

2. René Carmille, mort au camp de Dachau en 1945, a créé sous l'Occupation le Service national des statistiques (SNS) qui préfigurait l'Insee, et est à l'origine du numéro d'identification unique (numéro de Sécurité sociale), encore aujourd'hui en vigueur.

déterminées, mais sans que les résultats soient approfondis, commentés et utilisés dans les pratiques publiques visant à déterminer les choix. La place respective du sentiment et de l'imaginaire, d'une part, des analyses fondées sur des rationalités, d'autre part, dans les discours politiques, avantage nettement les premiers aux dépens des seconds.

Il en va autrement depuis le dernier conflit mondial pour de multiples raisons, tenant aux contraintes du moment (par exemple la gestion des crédits du plan Marshall), à la forte intervention de l'État dans l'économie et dans le « social » pendant et après la guerre, à la place que tiennent les nouvelles générations formées au calcul économique apparues peu avant la guerre (Paul Reynaud à droite, Pierre Mendès France à gauche), au développement des instruments administratifs aptes à la production statistique (l'Insee après le SNS), enfin, sans doute, à l'essor de l'éducation de masse. Des indicateurs chiffrés deviennent d'emploi courant, comme la mesure du coût de la vie (dans un temps d'inflation) ou l'évolution du salaire minimum. Le plan de modernisation et d'équipement et son instrument financier (le FME puis FDES³) sont à la fois le cadre pédagogique et l'outil de mise en œuvre d'une approche quantitative bientôt modélisée de l'action publique. La Sécurité sociale et ses « risques », le contrôle des prix, les investissements dans les infrastructures (reconstructions) requièrent des approches de même nature. Il a fallu à la fois développer les recueils de données utiles et les diffuser. La dernière moitié du xx^e siècle se déroule donc comme une lente progression de l'univers statistique, tant dans la fabrication que la diffusion, dont la loi du 7 juin 1951 est une expression majeure. Pour ne prendre que deux exemples de ces développements : on voit, d'une part, se manifester les souhaits des milieux patronaux de disposer de données leur permettant d'analyser les marchés ; on constate, d'autre part, que

3. Respectivement Fonds de modernisation et d'équipement et Fonds de développement économique et social, comptes du Trésor où s'inscrivaient, en loi de finances, les crédits d'investissement découlant du plan.

l'apparition de nouvelles politiques publiques s'accompagne (parfois avec un décalage chronologique) de la confection des indicateurs indispensables à leur définition (on pense ici, en guise d'illustration, à la politique de la ville de la fin du siècle⁴). Les instruments de diffusion se multiplient, désormais bien établis (*Tableaux de l'économie française...*), auxquels le « Net » a donné une ampleur supplémentaire. Il en va de même dans l'ensemble des pays industrialisés.

Le nombre, de répulsif, devient un instrument de vérité et, comme on dit aujourd'hui, de crédibilité. Il est inséparable de la connaissance : « *L'unique argument qui frappe les gens modernes – les chiffres* », note Hannah Arendt comme une évidence⁵. Leur fondement n'est plus contesté ; seulement, par des échanges de nature épistémologique, leur pertinence, leur cohérence, les commentaires dont ils sont assortis. Le développement des services statistiques ministériels, en particulier dans des départements qui y ont été longtemps rétifs (on pense au ministère de l'Intérieur), en est une autre expression.

Remise en cause de l'expertise

Diverses évolutions de nos sociétés développées « post-modernes » remettent en cause ce développement certes inégal selon les moments, mais incontestable sur le long terme. Mettons de côté, comme déjà identifiés, d'autres facteurs comme le rétrécissement du périmètre d'intervention de la puissance publique, qui réduit le besoin en données d'intérêt général, pour nous concentrer sur ce qui s'applique autant à la connaissance elle-même qu'aux instruments qu'elle entend employer. Sans prétendre ici à l'exhaustivité, on voudrait rendre le lecteur sensible à quelques considérations.

Et d'abord l'évolution de la demande.

4. Cf. Dans *Insee, Actualités Magazine*, mars 2007, le dossier « Statistiques pour la politique de la ville » (pages 2 à 8). Voir aussi les journées thématiques du Conseil national de l'information statistique.

5. *Nous autres réfugiés*, Paris, Allia éditions, 2019, p. 18.

Pour autant qu'on puisse la caractériser en termes simples, il est possible de penser qu'aujourd'hui (pas nécessairement demain) la demande de données évolue. Moins soucieuse de l'agrégation de données inchangées dans un cadre national, elle est tournée vers des données attachées à un horizon nettement plus individuel (données démographiques ou économiques de ma ville...) ou à un univers transnational dans des domaines déterminés (nombre d'étudiants étudiant la médecine à l'étranger, rapidité de la fonte des glaciers...), domaines eux-mêmes soumis à de fortes variations temporelles. Il en va des quêtes d'informations comme des aspirations et des revendications sociales : elles s'expriment moins qu'auparavant dans des cadres institutionnels définis et invariables.

On peut estimer aussi que le principe même de données robustes soigneusement travaillées fait moins recette. Les médias dominants, y compris ceux réputés sérieux hier, accordent une place croissante au for intérieur, aux effets des événements sur les sentiments et aux émotions des personnes. Une telle approche est souvent jugée davantage « à dimension humaine », « proche des gens », conforme « à la vraie vie », selon un vocabulaire employé *ad nauseam*. Naturellement, en parallèle, est instruit le procès des « experts » de tout acabit, notamment les producteurs de nombres peu sensibles à l'émotion. À ceux qui feraient valoir que ces accusations ont toujours existé, on demandera de prêter attention non à cette seule existence, mais à la tendance dominante de la parole publique, entendue dans un sens très général. Il est facile de montrer aussi le caractère inédit de la remise en cause de l'expert⁶ comme un sachant privilégié, au motif que ce qu'offre le marché permet d'affirmer que le point de vue qu'il défend n'est qu'une opinion parmi d'autres et que celle que l'on peut se forger soi-même par ailleurs pèse d'un poids équivalent, et même moins discutable puisque moins attachée à des intérêts (de système, d'appareil, d'habitudes intellectuelles...).

6. Terme désormais très amphibologique.

Il faut aussi prêter attention aux modèles qui font recette parmi les jeunes générations. À cet égard, les inquiétudes sur le peu d'appétence pour la connaissance (dont témoignent les enseignants des collèges), sur les difficultés d'attention, sur les facilités de dispersion, doivent être prises avec sérieux, même si ces phénomènes bien réels sont compensés par d'autres acquisitions relatives au raisonnement et à l'appréhension technique. Là encore, il convient de discerner quelle est la tendance dominante parmi le plus grand nombre.

Des sources en abondance

Du côté de l'offre à présent, il faut rappeler deux phénomènes que chacun a naturellement à l'esprit.

Le premier d'entre eux est la prolifération des nombres.

Il va de soi que, depuis longtemps, les sources se diversifient et les productions s'accroissent. Le volume de la production statistique doit être corrélé avec des notions aussi diverses que le développement économique national, celui des échanges au niveau international, la capacité de professions à imposer à leurs membres les sujétions inhérentes à la collecte de données⁷, les outils techniques permettant les calculs. Notre époque a donc vu s'accroître le nombre des statisticiens et des statistiques. Mais la profusion de ces dernières n'a jamais été aussi considérable et, plus encore, leur accessibilité n'a jamais été aussi aisée. Est en cause naturellement le développement du Web et ce qu'il autorise comme accès, « en un clic », à un nombre d'informations sans limite mesurable.

Ainsi, une consultation du site du Bureau gouvernemental australien des statistiques m'apprend qu'au recensement de 2016, 52 % des habitants se déclaraient croyants d'une religion chrétienne, 2,6 % se réclamaient de l'islam et 2,4 % du bouddhisme, 30 % se disant sans

7. Par exemple, les publicitaires qui ont imposé la mesure des « audiences » de médias.

religion⁸. Constat doublement intéressant : en ce qu'il révèle la déclaration que font les résidents de ce pays de leur croyance, mais aussi en ce qu'il fait état d'une donnée dont l'utilisation en France est sévèrement encadrée, la conviction religieuse relevant du seul domaine de l'intime⁹. Ou bien l'édition numérique du *Guardian* me fait savoir en juillet 2018 qu'il s'est produit 15 973 homicides au Mexique du 1^{er} janvier au 30 juin cette année-là, soit le nombre le plus élevé depuis 1997, date à laquelle on a commencé de répertorier cette donnée¹⁰. Là encore, la leçon est double : sur le taux d'homicides dans le pays et sur la date à laquelle le *Secretario Ejecutivo del Sistema Nacional de Seguridad Publica* a estimé l'information digne d'être publiée.

On voit les vertiges qu'ouvrent ces abîmes de savoirs : vertige de la connaissance à satisfaire sur tout et n'importe quoi ; vertige du démiurge qui, à partir de ces données multiples, peut opérer lui-même les combinaisons qu'il veut pour parvenir à un savoir qui lui sera propre ; vertige enfin du scientifique qui se demande quelle valeur il convient de conférer à chacun de ces nombres ainsi offerts au public.

Car tous ne se valent pas. Mais si, au plan national, la caution de fiabilité d'un nombre ressort des compétences, des prérogatives et des distinctions dont sont investis des producteurs, et non d'autres (cf. les débats récurrents sur le nombre de participants d'une manifestation sur la voie publique), comment opérer une telle distinction dans l'ensemble d'informations qui proviennent d'origines multiples et inconnues ? Un statisticien justifie son point d'aboutissement (cf. les éléments méthodologiques qui figurent toujours dans *Économie et*

8. <http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by+Subject/2071.0~2016~Main+Features~Religion+Data+Summary~70>

9. Selon l'article 8 de la loi informatique et libertés : « Il est interdit de traiter des données à caractère personnel qui révèlent... les convictions religieuses ou philosophiques... d'une personne physique. »

10. www.theguardian/world/2018/jul/23/mexico-crime-homicides-violence-up-report

*Statistique*¹¹ – notamment dans les encadrés et la bibliographie terminale – ou les revues similaires) ou peut être sommé de le faire. Les nombres publiés... en nombre se passent désormais volontiers de ces bagages : leurs conditions d'élaboration, par conséquent la mesure de leur véracité, demeurent souvent inconnues. Reste la magie du signe et la tromperie qui fait ressembler comme deux gouttes d'eau un mensonge éhonté et une vérité établie, et plus encore, si l'on peut ainsi s'exprimer, sur « la toile ». L'usage de ces nombres n'en est en rien dissuadé. Les *fake news* – les fallacies – ne sont en rien une invention ; ce sont leur volume et leur apparence qui le sont. Et l'on voit dans les médias apparaître les articles de « décodage » qui peuvent valoir aux statisticiens de rendre de fiers services, mais seulement au second degré et sur des terrains qu'ils n'ont pas nécessairement choisis ; ainsi que des « journalistes de données » qui se font statisticiens, par l'utilisation de nombres en libre accès, comme l'a montré un « Café de la statistique ».

Le Web ne se contente pas de rapporter des nombres : il en est lui-même producteur, dans le nombre de connexions ou de « clics » désapprouvateurs ou non. Les sociologues peuvent se réjouir de trouver dans de nombreux « sites » la liste des articles (ou assimilés) les plus regardés (ce qui ne signifie pas lus) ou des commentaires les plus « suivis » : ce sont, après tout, des indices de comportement.

La fin des précautions ?

Le second phénomène réside dans la prolifération des données statistiques, résultant elle-même de la multiplication des sources.

Les comportements sociaux (démographiques, familiaux...) ou l'activité économique ont eu pendant longtemps des nombres limités

11. À titre d'exemple, les développements de la page 121 et les tableaux de l'article de Jean Barthélémy, Vincent Bignon et Benoît Nguyen, « Politique monétaire, collatéral illiquide et crédits à l'économie pendant la crise européenne de la dette souveraine », *Économie et Statistique, Economics and Statistics*, 2017, n° 494-495-496, p. 119-140. Voir aussi, de manière plus générale, la rubrique « Définitions, méthodes et qualité » sur le site Web de l'Insee.

de séries de données faisant l'objet de recueil, séries elles-mêmes définies par un nombre limité de statisticiens : l'enjeu était de parvenir à une information synthétique et, à cette fin, de collecter les données élémentaires en rapport avec l'information voulue. Désormais, ce n'est pas que « tout fasse sens » qui pose problème, mais que « tout peut être recueilli », d'une part, et de manière massive, d'autre part, sans qu'une intention de départ formalisée en précise l'usage et l'utilité. Au surplus, ces états ne sont pas seulement collectés, mais mis en parallèle avec d'autres séries distinctes voire sans rapport avec les originaux. Au lieu que la collecte fasse l'objet de démarches particulières (constitution d'un objet, utilisation d'une source et recueil de ce qui s'y trouve par des moyens très variés¹², avant exploitation mathématique, mise en ordre de résultats et analyses), elle se fait automatiquement, sans intervention humaine, mais par connexion, et de manière continue. Ainsi se constituent les Big Data, destinés à mettre en lumière, par le biais d'algorithmes sur lesquels on va revenir, des fréquences et, grâce au rapprochement de plusieurs de ces dernières, des causalités.

Dans ces amas de données, des comparaisons peuvent s'opérer selon des hypothèses formulées *ex ante* : les pratiquants de sport réguliers ont moins d'accidents vasculaires cérébraux que les autres ; on va donc croiser les données recueillies sur les *joggers* et celles issues des personnes hospitalisées pour cette affection, en nombre tel que la robustesse du résultat sera élevée. Mais les Big Data peuvent être (et sont souvent) utilisées sans la moindre hypothèse de départ : l'offreur de services sur Internet n'est en rien statisticien ; il a cependant à sa disposition des quantités très grandes de données, d'ailleurs à grande

12. Cf. l'évolution des méthodes du recensement de la population française ; voir notamment Jean-Noël Biraben, « Inventaire des listes nominatives de recensement en France », *Population*, 1963, 18-2, p. 305-328 ; Insee, département de la démographie, « Présentation du recensement rénové », *Population*, 2000, 55 4-5, p. 803-811 ; Alain Godinot, « Pour comprendre le recensement de la population », *Insee Méthodes* hors-série, mai 2005 ; Luc Albarello, Étienne Bourgeois et Jean-Luc Guyot, *Statistique descriptive*, de Boeck Supérieur, 2010.

valeur commerciale, qu'il peut utiliser à sa guise, sans précautions particulières, pour en tirer des résultats inédits, et en inférer des causalités plus ou moins assurées, plus ou moins simplistes, en omettant les écheveaux de phénomènes complexes. Plus prosaïquement, le nombre de « faux vrais », marginal dans un échantillon, s'accroît évidemment beaucoup dans de tels ensembles numériques.

Ces données peuvent enfin être l'objet, comme on le sait, d'échanges commerciaux et être utilisées par plusieurs personnes successivement, à des fins très variées. Ce qu'elles expriment (le nombre de pas de quelques millions de promeneurs du dimanche) peut donc, au fil des utilisations, n'avoir que de lointains rapports scientifiques avec l'investigation initiale : en somme, l'information, pressurée, essorée, est plus riche, au risque de déformation ou de dénaturation. La précaution épistémologique de la statistique n'est plus de mise, même si on ne doit pas douter que les Big Data apporteront des vérités considérables.

Secret, indépendance et transparence

Reste enfin à marquer l'évolution du « secret statistique ».

La problématique du secret statistique¹³, défini dans la loi du 7 juin 1951, évolue peu favorablement.

Tout ne se publie pas, mais tout ce qui n'est pas public a des raisons (qu'on peut trouver trop étroites ou trop larges) de ne pas l'être. Ces motifs, qui sont ceux de la loi informatique et libertés du 6 janvier 1978 comme du règlement général de protection des données 2016/679 de l'Union européenne (RGPD), s'inspirent de la nécessaire protection de la vie privée. L'article 6 de la loi de 1951 s'énonce ainsi : « Les renseignements individuels figurant dans les questionnaires... et ayant trait à la vie personnelle et familiale et, d'une manière générale, aux faits et comportements d'ordre privé ne peuvent... faire

13. Voir Insee, *Guide du secret professionnel*, version du 27 juillet 2018, 9 p. ; ainsi que *Le Secret statistique*, de Jean-Pierre Le Gléau, paru aux éditions EDP-Sciences, 2019.

l'objet d'aucune communication de la part du service dépositaire... » Sont également soumis à des restrictions (de moindre durée) les renseignements économiques et financiers individuels, pour des motifs également compréhensibles, de même que des nombres relatifs aux secrets protégés par la loi (secret défense...) ou aux activités qui doivent rester confidentielles (secret des affaires).

Ces restrictions emportent *a contrario* l'idée que toutes les autres informations – tous les autres nombres – sont publiables.

Cette obligation du secret dans certaines hypothèses est inséparable de l'obligation de transparence dans toutes les hypothèses. Les utilisateurs de nombres doivent savoir d'où ils proviennent et comment ils sont établis. Ce n'est pas là seulement une question technique, mais une contrainte qui relève de notre vie démocratique. Les dictatures n'aiment que les « bons » chiffres, et pas seulement ceux des « élections ». La démocratie s'oblige¹⁴ à supporter les mauvais, s'ils sont incontestables. Pouvoir décrire rigoureusement la manière dont les données sont produites conduit en parallèle à assurer l'indépendance du statisticien qui, à son tour, est un élément central de la fiabilité des nombres.

Les personnes privées détentrices de données (auxquelles la loi de 1951 n'est pas applicable) en sont propriétaires et en ont donc le libre usage, dans le cadre des dispositions de la loi informatique et libertés. Toutefois, il est très souhaitable que les nombres qu'elles établissent soient soumis aux mêmes règles de transparence. Mon transporteur, par le truchement de la puce de ma carte d'abonnement, mon banquier, *via* ma carte de crédit, mon opérateur téléphonique, savent ce que je dépense, où je me déplace et quelles sont mes relations. Toute règle d'anonymat disparaît. Une étude américaine a montré que l'utilisateur d'une carte bancaire pouvait être identifié après trois achats avec 80 % de chances et, après quatre, avec une

14. De bonne ou de mauvaise grâce, naturellement. Vue du ministère de l'Économie, comme l'auteur a pu le constater, l'indépendance de la statistique est parfois mal comprise et même mal supportée.

Mesure et perception de la mesure
 (10.1), 183
 Mobilité (4.3), 94

O

Opinion publique (5.3, 14.1), 107,
 232

P

Pauvreté (2.3), 59
 PIB (9.4), 177
 Population mondiale (1.1), 35
 Prélèvements obligatoires (9.2), 170
 Profilage (12.1, 2.1), 207, 53
 Progrès des sociétés (9.4), 177
 Projection (11.1), 197
 Prospective (11.1), 197

R

Réchauffement climatique (13.1),
 217

Recherche-développement (4.1), 87
 Retraite (5.1), 101
 Réussite scolaire (9.3), 174

S

Santé (8.1), 146
 Sécurité routière (3.3), 76
 Statistiques communautaires (9.1),
 167
 Statistiques ethniques (7.2), 130
 Suicide (7.5), 140

T

Tourisme (3.4), 79
 Transition démographique (1.1), 35
 Travail (8.3), 153

V

Valeur ajoutée (13.3), 223

LE NOMBRE ET LA CITÉ

La statistique éclaire-t-elle les questions de société ?

Pourra-t-on nourrir dix milliards d'humains ? Faut-il encore construire des logements ? Quelle est l'ampleur de l'immigration ? Comment la richesse est-elle partagée ? La statistique se propose d'éclairer nombre de questions aussi déterminantes pour l'avenir de nos sociétés. Y parvient-elle ?

Ce livre est né d'une expérience peu ordinaire. Depuis 2005, plus de cent vingt débats publics ont confronté la parole d'experts à celle de citoyens « comme vous et moi », animés du désir de mieux comprendre. Le produit de ces « Cafés de la statistique » forme un corpus de connaissance dont les principaux enseignements méritent d'être diffusés.

La statistique irrigue la vie de la cité parce qu'elle aide à comprendre le monde, à rationaliser et à évaluer les politiques publiques, servant ainsi la démocratie. Elle trouve ses limites quand elle doit surmonter les barrières culturelles ou quand ses productions font l'objet d'utilisations tendancieuses. Ces risques lui imposent d'être explicite sur ses conventions, ses pratiques et ses méthodes en respectant une déontologie sans faille.

Le lecteur pourra se laisser conduire par le plan de l'ouvrage ; ou bien il butinera d'un sujet à l'autre, au gré de ses inclinations. Sa lecture l'étonnera parfois ; elle l'éclairera souvent. Les auteurs prennent ici soin de s'adresser à des non spécialistes.

Pont-Neuf, ce sont sept auteurs, membres de la Société française de statistique, qui ont participé à l'organisation et à l'animation des Cafés de la statistique et rédigé en commun cet ouvrage : Brigitte Belloc, Alain Godinot, René Padieu, Sandra Roger, Jean-François Royer, Marion Selz et Olivier Vasseur.

Prix : 22 €

ISBN : 978-2-7598-2374-1



9 782759 823741

La Société Française de Statistique (SFdS), association reconnue d'utilité publique, a pour objectif de favoriser les développements de la statistique et d'assurer la représentation de l'ensemble des utilisateurs, enseignants et chercheurs dans ce domaine. Sa collection « Le monde des données » propose des ouvrages accessibles à un très large public. La SFdS témoigne ainsi de sa volonté de s'ouvrir davantage vers la société civile.

SFdS

edp sciences
www.edpsciences.org