



# Les classes de risques vont-elles plus loin que les stéréotypes ?

## Modèles prédictifs, entre explications fallacieuses et clichés ?

### Généralisation, stéréotypes et café du commerce

Dans *Système 1 / Système 2* : Les deux vitesses de la pensée, Daniel Kahneman revient longuement sur l'importance des stéréotypes pour comprendre bon nombre de prises de décision. Le système 1 est utilisé pour la prise de décision rapide : il nous permet de reconnaître les gens et les objets, nous aide à orienter notre attention, et nous encourage à craindre les araignées. Il est basé sur des connaissances stockées en mémoire et accessibles sans intention, et sans effort. On peut l'opposer au système 2, qui permet une prise de décision plus complexe, exigeant de la discipline et une réflexion séquentielle. Dans le premier cas, notre cerveau utilise les stéréotypes qui régissent les jugements de représentativité, et utilise cette heuristique pour prendre des décisions. Si je cuisine un poisson pour des amis venus manger, j'ouvre une bouteille de vin blanc. Le cliché « le poisson se marie bien avec le vin blanc » me permet de prendre une décision rapidement, sans avoir à réfléchir. Les stéréotypes sont des affirmations concernant un groupe qui sont acceptés (du moins provisoirement) comme des faits concernant chaque membre. Qu'ils soient corrects ou justes, les stéréotypes sont l'outil de base pour penser les catégories dans le Système 1. Mais bien souvent, une réflexion plus poussée, plus construite – correspondant au Système 2 – permettra de prendre une décision plus judicieuse, voire optimale. Sans choisir n'importe quel vin rouge, un pinot noir pourrait peut-être aussi parfaitement convenir aux rougets grillés.

En France, des clichés comme « les universités sont

des fabriques de chômeurs » ou « l'université est un choix par défaut » sont des stéréotypes qui ont fait du tort pendant des dizaines d'années. « *To generalize is to be an idiot, to particularize is the alone distinction of merit* », écrivait William Blake vers 1800, annotant un discours du peintre Joshua Reynolds. Stigmatiser toute une population à cause d'une minorité, dans un processus de prise de décision correspond à une généralisation fallacieuse, souvent punie par la société. Punition morale, mais parfois aussi légale (dans les embauches par exemples), dans une société qui tend à être civilisée, demandant de ne pas tirer des conclusions erronées sur un individu à partir des statistiques d'un groupe auquel il se rattacherait. Mais n'est-ce pas ce que fait l'actuaire tous les jours ?

### The usual suspects

Pour Schauer (2009), cette « généralisation », condamnée par William Blake, est probablement la raison d'être de l'actuaire : « *to be an actuary is to be a specialist in generalization, and actuaries engage in a form of decision-making that is sometimes called actuarial* ». Si je décide d'assurer une voiture de sport, je me vois prêter des caractéristiques de conduite à risque qu'ont probablement la majorité des propriétaires de voitures sportives, attributs que je ne partage pas peut-être pas. Et comme nous le notions dans l'introduction, les compagnies d'assurance, bien sûr, ne sont pas les seules qui fonctionnent de façon actuarielle pour reprendre la définition de Schauer (2009). Nous le faisons tous, beaucoup plus souvent que la plupart d'entre nous ne le reconnaissent. Nous

le faisons lorsque nous choisissons les compagnies aériennes sur la base de leurs antécédents en matière de sécurité, de leur ponctualité ou de la perte de bagages. Nous opérons ainsi lorsque nous associons des caractéristiques personnelles (un tatouage visible, des vêtements noirs ou au contraire très colorés) à des caractéristiques comportementales (comme une propension à la violence) que ces caractéristiques personnelles sembleraient indiquer. Et nous fonctionnons ainsi lorsque nous nous livrons à des stéréotypes en fonction de la nationalité, par exemple en qualifiant les Italiens de personnes très expansives, ou les Écossais de radins, tout en reconnaissant en même temps que des stéréotypes plus pernicioeux en fonction de l'origine ethnique, du sexe, de l'orientation sexuelle sont trop répandus aujourd'hui ! Comme l'indique la mauvaise connotation du mot « préjugé », beaucoup de gens croient qu'il est injuste de prendre des décisions individuelles sur la base de caractéristiques de groupe non universelles. Même si les attributions de groupe ont une base statistique solide. Car la grande différence entre l'actuariat et la vie quotidienne est que les actuaires se doivent d'utiliser un grand nombre d'observations. À titre personnel, je peux ainsi décider de ne plus voyager avec telle compagnie aérienne parce que sur trois voyages, j'ai eu deux fois des ennuis. Avant de décider qu'une assurance-voyage mérite une prime plus élevée lorsque le vol s'effectue avec cette compagnie, il faut plus que trois observations ! En fait, la question est souvent de savoir si le refus d'une compagnie d'assurance de fournir une couverture, ou d'augmenter les primes qu'elle demande pour la même couverture, constitue une injustice lorsqu'elle se fonde sur une généralisation actuariellement justifiée (mais peut-être non universelle). Comme le notait Leemans (2000), la question avait ainsi été posée au législateur lorsque des assureurs avaient observés que les femmes juives originaires d'Europe orientale étaient particulièrement vulnérables au cancer du sein et de l'ovaire. Fin 2012, la Cour de Justice Européenne mettait fin à toute discrimination basée sur le genre des assurés : plus question pour les assureurs de différencier les tarifs des produits d'assurance selon si l'adhérent est un homme ou une femme. Mais l'utilisation de l'âge reste autorisée. En effet, l'âge est souvent un indicateur d'une possible diminution de la vision ou de l'audition, du ralentissement du temps de réaction (et du risque accru d'incapacité soudaine), etc. Et, bien qu'il y ait beaucoup de variations individuelles,

les données disponibles fournissent une justification empirique importante.

### Machines, causalité, et stéréotypes

Une grande critique des modèles de « machine learning » est le manque d'interprétation. Mais bien souvent, la validation de modèles économétriques passe par une narration construite autour de stéréotypes. Et cette narration est essentielle, comme le rappelle Pearl & Mackenzie (2018). En effet, dans l'échelle de la causalité – « The Ladder of Causation » – on distingue trois niveaux. Au premier niveau, on retrouve la notion d'association (ou de corrélation), voire de probabilité conditionnelle, qui servent de base à la constitution de stéréotypes : si on observe

$P[\text{carries} \mid \text{se brosser les dents}] < P[\text{carries} \mid \text{ne pas se brosser les dents}]$

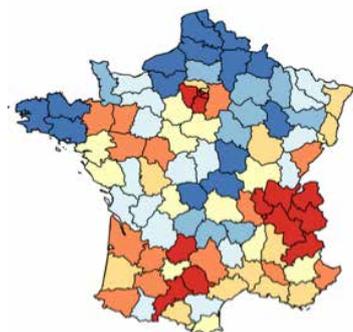
on associera le fait de se brosser les dents avec la baisse de la probabilité d'avoir des caries. C'est la base aussi des méthodes de régression, qui reposent sur des corrélations entre la variable d'intérêt et d'autres, abusivement appelées explicatives.

Le premier niveau de l'échelle permet de répondre à la question « *what if I see... ?* » (par exemple « combien de caries peut-on espérer chez une personne qui se brosse les dents tous les deux jours ? ») et cette tâche peut être parfaitement accomplie par une machine. Le second niveau est celui qui permet de comprendre un effet, une intervention (une politique publique). La question est alors « *what if I do... ?* ». Pour reprendre notre exemple, que se passerait-il si on forçait la personne à se brosser les dents tous les jours ? On voit



**Arthur Charpentier**  
Professeur, Université du Québec à Montréal  
et éditeur de <http://freakonometrics.hypotheses.org>

Espérance de vie



Licence de Tennis



Agence de publicité

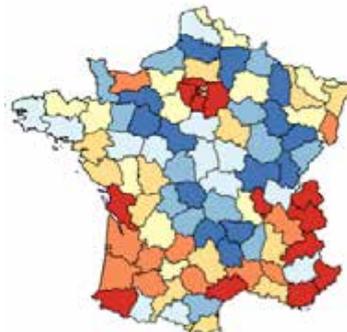


Figure 2 : Espérance de vie à la naissance pour les hommes, à gauche. Au centre, nombre de licences de tennis pour 100 000 habitants (source FFT). À droite, nombre d'agences de publicité pour 100 000 habitants. Source : INSEE, code NAF 7311Z). Visualisation de l'auteur.

d'ailleurs arriver le troisième niveau de l'échelle, posant la question « *what if I had done...?* », et reposant sur l'idée de modèle contrefactuel. On ne se contente plus de mesurer des corrélations, on va construire un modèle expliquant ce qui se passerait en faisant un changement dans les variables causales : que se passerait-il vraiment si l'enfant qui ne se brossait pas les dents commençait à le faire ? Pour Pearl & Mackenzie (2018) un être humain (voire un actuair) peut faire ces raisonnements poussés qu'une machine ne peut pas (encore) faire. Et bien souvent, ces schémas causaux passent par des stéréotypes. Comme le rappelle Charpentier & Diago Barry (2015), en épidémiologie, les chercheurs se sont longtemps interrogés sur l'explication à donner au fait que les petits bébés de femmes fumeuses avaient des probabilités de survie plus importantes que les bébés de mères non fumeuse. L'intuition que quelque chose ne va pas vient bien de préjugés, de stéréotypes que nous avons, et que ne peut avoir une machine.

### Quand les actuaires se racontent des histoires

Comme le notait Antonio & Charpentier (2017), la « gender directive » a dérouter nombre d'assureurs qui utilisaient le genre pour construire leur tarif, car cette dernière était très corrélée avec la fréquence de sinistres. Mais en introduisant des données télématiques, le genre n'était alors plus significatif dans la régression. Le genre a ainsi longtemps servi de proxy pour capturer un effet que l'on peut observer à l'aide de données télématiques, donnant lieu à bon

nombre de clichés sexistes et autres stéréotypes. Mais les histoires permettent aussi de trancher entre une corrélation fallacieuse (« spurious correlation ») et une corrélation qui pourrait être interprétée. Sur la Figure 2, on a l'espérance de vie à la naissance, variable que l'on pourrait chercher à expliquer dans un contexte d'étude de rentes, par exemple, par département français. À droite, deux variables prises au hasard : le nombre de licenciés d'un club de tennis, et le nombre d'agences de publicités. Les stéréotypes, c'est ce qui va nous permettre de construire un graphique causal, permettant de comprendre pourquoi la corrélation est aussi forte entre ces variables et l'espérance de vie.

### L'hyper-individualisation comme réponse ?

Si William Blake condamnait les stéréotypes en affirmant « *to generalize is to be an idiot* », il allait aussi clairement plus loin, poursuivant avec « *to particularize is the alone distinction of merit* ». Cette individualisation est d'ailleurs prônée par de plus en plus d'assureurs, voire souhaitée par de nombreux assurés. Mais comme le rappelait Grace & Terry (2002), de nombreux assurés souffrent d'un important biais d'optimisme – « si j'ai un accident, ce n'est pas ma faute » – les poussant à douter de la classification faite par l'assureur – « je suis moins risqué qu'eux autres ». Est-ce pour autant injuste ? Une société juste est-elle nécessairement une société dans laquelle chaque individu est traité comme un ensemble isolé d'attributs uniques, exigeant une attention individualisée ?

REFERENCES : Antonio, K. & Charpentier, A. (2017). La tarification par genre en assurance, corrélation ou causalité?. Risques, 110 : 107-110./// Charpentier, A. & Diago Barry, A. (2015). Big data : passer d'une analyse de corrélation à une interprétation causale. Risques, 101 : 107-111./// Grace, J. & Terry, M. (2002). Exploring the Causes of Comparative Optimism. Psychologica Belgica. 42: 65-98 /// Kahneman, D. (2011). Système 1 / Système 2 : Les deux vitesses de la pensée, Flammarion./// Leemens, T. (2000). Selective Justice, Genetic Discrimination, and Insurance: Should We Single Out Genes in Our Laws? McGill law journal. Revue de droit de McGill 45(2):347-412./// Pearl, J. & Mackenzie, D. (2018). The Book of Why: The New Science of Cause and Effect. Basic Books./// Schauer, F.F. (2009). Profiles, Probabilities, and Stereotypes. Harvard University Press.