



Terroirs de whisky

PATRICK RIVES

L'analyse statistique des données montre comment déguster les scotches.

L'Écosse, terre d'élection du whisky ! Il s'y nomme scotch, et il en existe plusieurs appellations : les *single malt* sont obtenus à partir d'orge fermentée provenant d'une seule distillerie ; les *blended* sont des mélanges de liqueurs de différentes qualités, pouvant être issues de plusieurs régions de production. Évidemment les *single malt* sont les préférés des connaisseurs, qui scrutent leurs qualités : nez, couleur, corps, bouche et finale.

La provenance d'un *single malt* détermine-t-elle ses qualités organoleptiques ? Si oui, peut-on regrouper les scotches en terroirs d'où proviendraient des boissons analogues ? Afin de répondre à ces questions, François-Joseph Lapointe et Pierre Legendre, de l'Université de Montréal, ont analysé les données réunies par le dégustateur Michael Jackson, qui a décrit, goûté et jugé les quelque 330 *single malt* que produisent les 109 distilleries les plus prestigieuses d'Écosse.

Les techniques d'analyse des données utilisées extraient des informations relatives à la population des *single malt*. Ces derniers sont décrits par des caractères (la couleur, le nez, la bouche, le corps et la finale), qui peuvent prendre plusieurs modalités. Ainsi le nez peut être aromatique, tourbeux, léger, doux, frais, sec, fruité, herbeux, salé, sherry, épicé, riche ; ou encore le corps peut être moelleux, moyen, plein, rond, mielleux, léger, ferme, gras...

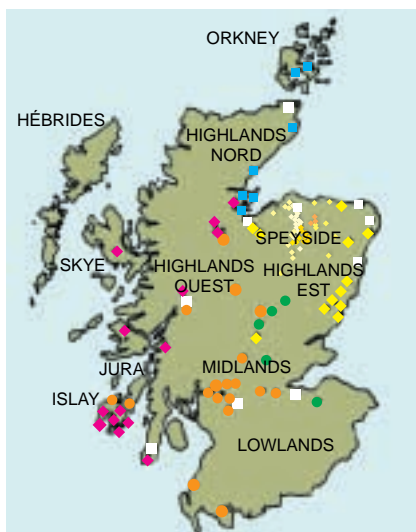
Comment découper la population des *single malt* en groupes d'individus analogues, afin de déterminer ensuite s'ils proviennent d'un même terroir ? Les statisticiens canadiens ont choisi de ne retenir qu'un scotch par distillerie, afin d'analyser ainsi 109 *single malt*.

Pour quantifier les impressions, ils les ont considérées chacune comme un caractère, qui prenait une valeur 1 quand il était présent dans un *single malt*, et 0 dans le cas inverse. Puis ils ont regroupé les données statistiques dans un tableau découpé en cases : en lignes, les individus ; en colonnes, les caractères.

Ils ont ensuite calculé une distance entre les *single malt* pris deux à deux en divisant le nombre de caractères communs aux deux whiskies comparés par le nombre

total de caractères relevés pour l'une ou l'autre des deux boissons. La distance est alors égale à 1 moins ce paramètre. L'ensemble de ces distances forme une nouvelle matrice, où la valeur d'une case représente la distance entre les whiskies respectivement associés à la ligne et à la colonne. Plus cette distance est petite, plus les whiskies sont proches.

Il restait alors à regrouper les individus. La classification hiérarchique que l'on utilise alors crée la partition progressivement, en agrégeant deux à deux les individus les plus proches : l'agrégat est alors considéré comme un individu nouveau, qui remplace les deux individus agrégés



Les whiskies gustativement analogues sont signalés par un même symbole.

(les modalités de ses caractères sont alors la moyenne de celles des deux individus). On continue ainsi à grouper jusqu'à ce que tous les individus soient rassemblés en une seule famille : on obtient un dendrogramme. Pour obtenir différentes partitions, on effectue une « coupure » du dendrogramme à une certaine distance de la racine. Plus on tranche près de celle-ci, moins les classes sont nombreuses.

Le dendrogramme obtenu pour les *single malt* se divise en deux branches : d'un côté, les whiskies plein or, corps sec et fumé ; de l'autre, des boissons ambrées, aromatiques, au corps léger, moelleux en bouche et de finale fruitée. Le premier groupe de 69 individus est alors subdivi-

visé en scotches ambrés, de corps charnu, fruité, gras et épicés, et les boissons dorées, au corps moelleux et léger, avec une finale herbeuse. Plus on s'éloigne de la racine, plus les groupes sont petits.

Revenons à la composante géographique. L'Écosse est divisée en trois régions de production de scotch : les Highlands, les Lowlands et l'île d'Islay. Ces trois régions sont découpées en treize districts. Les auteurs ont construit un nouveau tableau en mettant en ligne et en colonne les 109 scotches. L'intersection d'une ligne et d'une colonne contient un 0 si les scotches correspondants *i* et *j* sont issus du même district, elle est 1 dans le cas contraire.

Quand on compare cette matrice à celle qui a engendré le dendrogramme, on observe que le découpage des régions de production correspond à la partition, tirée du dendrogramme, en 6 ou en 12 groupes. Ces groupes correspondent à des terroirs.

Mieux encore, les analyses qui conduisent à un tel découpage prouvent que l'eau, le sol, le microclimat, la température et l'environnement en général sont les facteurs déterminant les caractéristiques des *single malt*. Les secrets et les traditions des distillateurs n'induisent que de petites différences.

Enfin, les statisticiens canadiens ont cherché les relations entre les cinq types de caractères. La réponse est nette : le nez, la couleur, le corps et la bouche ne sont pas des caractères indépendants. La couleur d'un scotch est reliée à son nez et à son corps. Le nez est lié à la bouche et, dans une moindre mesure, au corps. En revanche, la finale, c'est-à-dire l'impression qui reste en bouche quand on a fini de boire, ne dépend ni du nez, ni de la bouche, ni du corps, ni de la couleur, de sorte que la méthode de dégustation qui consiste à recracher la boisson après en avoir jugé la couleur, le nez, le corps et la bouche est critiquable, puisqu'elle néglige un paramètre fondamental.

Prochain rendez-vous *France Info* et *Pour la Science* le 30 novembre 1998, avec la chronique *Info Sciences* de Marie-Odile Monchicourt.