

Le sens de la mesure : nouveaux usages de la quantification

En 1995, réfléchir sur l'utilisation de la quantification en histoire revenait, pour Jean-Yves Grenier¹, à s'interroger sur les difficultés et les réticences des historien(ne)s face aux méthodes quantitatives – et sur la crise du quantitatif en histoire, après l'apogée des années 1960 et 1970. Son constat était toujours : l'usage des techniques quantitatives tendait alors à nier la spécificité même de l'activité historique – d'où finalement cette crise, renforcée par les critiques émises par les micro-historien(ne)s. Le plaidoyer en faveur de la quantification était avant tout un plaidoyer pour le recours aux « données quantitatives » – une expression qui paraît souvent aller de soi, alors qu'elle mêle les statistiques du passé et les sources, pas toujours chiffrées mais supposées faciles à mettre en série, produites par les interactions avec l'administration (actes d'état civil ou recensement des conscrits, par exemple). L'enthousiasme pour la quantification avait fait

oublier les biais de ces données et de certaines manières de les traiter, en particulier l'anachronisme qui menace la constitution de séries longues à partir de matériaux hétérogènes². Plus généralement, les procédés d'agrégation des données propres aux analyses labroussiennes d'histoire économique ou sociale semblaient inadaptés à des corpus historiques de plus en plus reconnus comme éclectiques et composites – et prenaient le risque de sacrifier la diversité des trajectoires individuelles à un mythique homme moyen. Enfin, les analyses affirmant mettre au jour des structures – notamment les analyses factorielles des correspondances, particulièrement développées en France, mais aussi les simples tableaux croisés de pourcentages – s'avéraient souvent statiques, peinant à prendre en compte les évolutions et les temporalités.

Face à ce constat, Jean-Yves Grenier dessinait des perspectives pour sortir de la crise. Il concluait en particulier son article sur la nécessité de considérer, dans le cadre du paradigme indiciaire de la micro-histoire, les chiffres (trouvés dans les sources ou produits par les historien(ne)s) comme des traces parmi d'autres, à interpréter à côté des textes et des tessons de poterie, sans leur conférer de dignité particulière ni les soupçonner plus que d'autres. Qu'en est-il plus de quinze ans plus tard ? Est-il prématuré de proclamer ou de prédire la sortie de crise de l'histoire quantitative ? Le temps est-il révolu où la valeur ajoutée d'une recherche historique se mesurait à l'aune des kilomètres de saisie ou de l'utilisation de procédures toujours plus obscures ? Nous aimerions ici donner une idée des nouveaux usages de la quantification en histoire. Les exemples qui jalonnent ce texte constituent délibérément un inventaire à la Prévert, mêlant travaux déjà publiés et en cours. Ils visent à offrir un aperçu de la diversité des approches, à donner envie de lire ces études, d'explorer, à leur suite, d'autres thèmes et de s'inventer d'autres sources.

Éclectisme des thèmes et des périodes : une formalisation sans chasses gardées

Les constats critiques des années 1990, portant sur une historiographie s'affirmant comme pionnière ou triomphante dans les années 1960 et 1970, mais encore bien présente dans la masse des productions historiennes des années 1980, semblent toujours valables. Néanmoins, leur cible s'éloignant dans le temps, est-il nécessaire de les rappeler ? Il nous faut – malheureusement, osera-t-on – répondre par l'affirmative si l'on en juge par la persistance de discours, émanant tant de « quantitativistes » que d'« anti-quantitativistes » autoproclamés, qui, aujourd'hui encore, n'imaginent d'histoire formalisée que labrousienne : préoccupée de grands objets, économiques ou sociaux, du peuple plutôt que des élites, de subsistances plutôt que de représentations, structuraliste, déterministe, peu ou prou marxiste – et réservée à un groupe de fanatiques, frottés de mathématiques et férus d'informatique. Cette cristallisation date en réalité non de Labrousse mais de Simiand, ou plutôt elle accompagne la réception longue de sa controverse avec Seignobos³. Elle se traduit en pratiques très concrètes, qui l'entretiennent à leur tour : cours d'initiation au quantitatif en histoire confiés à des doctorants en mathématiques ou informatique ; postes fléchés « quantitatifs » ou « numériques » à propos desquels la majorité du comité de sélection s'admet dès le début incompetente, déléguant l'appréciation des candidat(e)s à des « spécialistes », comme s'il ne s'agissait plus d'histoire ; chapitres « quantitatifs », dits aussi « méthodologiques » ou « sociologiques », faisant l'objet d'un isolement sanitaire au sein des thèses, parfois sans notes de bas de page, alors que le reste de la thèse évite scrupuleusement les tableaux et tout rappel des résultats « quantitatifs » en regard d'autres conclusions.

Tout cela, nous l'avons directement observé. Et pourtant, la situation a aussi changé. Nos propres parcours en témoignent⁴. L'histoire économique n'est pas au centre de nos travaux, l'échelle macro moins encore, puisque nos thèses résultent notamment d'une socialisation très micro-historienne⁵. Ce n'est pas par conviction idéologique que nous sommes venues aux méthodes quantitatives, mais par nécessité : nos sources nous y ont conduites, nos démonstrations s'y adossent, mais la quantification n'est pas pour nous une fin, plutôt un outil parmi d'autres. Et nous ne revendiquons en rien quelque étiquette de « quantitatistes ». Mais nous enseignons aussi les méthodes quantitatives depuis plusieurs années, dans des lieux divers et auprès d'apprenti(e)s chercheur(se)s d'universités aussi variées que leurs thèmes ou périodes de prédilection. Notre premier constat concerne l'augmentation, au fil des années, des effectifs d'un public ni prédestiné à quantifier ni cherchant à devenir quantitatif, mais tout simplement en quête d'aide.

C'est souvent le constat d'un volume de sources qu'il serait dommage ou impossible de traiter sans base de données, voire sans échantillonnage, qui conduit ces historien(ne)s à chercher une formation, ou plus souvent une assistance. Ainsi, face à la masse des dossiers de carrières d'employés de la poste dans l'Algérie coloniale (Annick Lacroix, ENS Cachan), des procès devant la justice de paix (Anaïs Albert, université Paris-1) ou encore des dossiers nominatifs de l'office HLM à Saint-Denis (Cédric David, université de Nanterre), ils-elles se demandent comment en tirer non seulement des exemples ponctuels ou une impression générale, mais bien des typologies raisonnées ou encore des conclusions claires – souvent pour évaluer la réalité d'idées reçues, comme l'association entre administration coloniale et traitement à part des « indigènes », entre consommations ouvrières et endettement ou encore entre immigrés et logement social.

Ces exemples pourraient être placés sous l'étiquette, large parce qu'imprécise, d'histoire sociale, et portent sur le *xx^e* siècle. Mais ni le volume des sources quantifiables – si l'on se déprend d'une appréhension routinière de celles-ci, tournant autour de l'âge, du salaire ou de la production – ni la volonté de prouver ne sont l'apanage d'un thème, d'un terrain ou d'une période. Pour parler de sources abondantes, évoquons l'une des matières premières privilégiées de l'histoire : les textes. Les analyses de discours, dont les pionniers avaient été trop rapidement oubliés⁶, sont souvent la technique qui semble la plus pertinente à de jeunes historien(ne)s sans appétence particulière pour la quantification, mais qui constatent la complémentarité d'une lecture éloignée et systématique, assistée par l'ordinateur, avec une lecture rapprochée et interprétative. Le travail de Sophie Cinquin sur les œuvres dédiées à Marguerite de Navarre en donne un bel exemple⁷. Là comme ailleurs, l'enjeu est finalement de banaliser la technique, ou plutôt les nombreuses techniques différentes qui peuvent permettre d'explorer systématiquement des textes, d'en faire un outil parmi d'autres plutôt qu'un étendard. Cette étude souligne aussi que l'histoire quantifiée des arts et de la littérature n'est pas condamnée à une approche externaliste, celle des marchés, carrières ou publics, ou encore de la construction des canons et des réputations, mais peut aussi être féconde lorsqu'elle touche le contenu et le style même des œuvres. L'étude des genres, mais aussi des sujets précis des peintures exposées – ou refusées par le jury – pour le Salon de 1827 menée par Eva Bouillo va jusqu'à l'examen visuel de tableaux qui peuvent être caractérisés de banals ou d'exceptionnels. L'analyse permet ainsi de discuter à la fois les choix de motifs des peintres, ceux de jurés soucieux de diversifier les œuvres exposées, voire le goût du public. Ces variations d'échelles et de questionnements possibles dès lors que l'on considère

pleinement l'œuvre comme une source peuvent tout aussi bien s'appliquer en histoire de la musique⁸.

Dans un texte à tous égards pionnier, Daniel Milo soulignait dès 1987 que l'histoire culturelle (entendue comme celle des pratiques culturelles des élites) avait sans doute, en réalité, bien plus de données aisément quantifiables à sa disposition que l'histoire matérielle du peuple⁹. Bien des travaux lui ont donné raison, tout en soulignant avec force que la formalisation pouvait s'appliquer à toutes les échelles. Même lorsque peu de sources sont conservées, la finesse avec laquelle on les lit produit des données riches, parfois avec peu d'individus (ou d'objets, de textes, etc.) mais tant de variables pour les décrire qu'une quantification s'avère heuristique, voire indispensable. Ils ont aussi montré que, contrairement à ce que craignait encore Milo, il n'est pas en soi plus difficile de quantifier lorsqu'on s'éloigne du présent, du monde des statistiques étatiques et des sondages d'opinion. La perte de confiance dans ces derniers, souvent traités auparavant comme des sources privilégiées car plus « fiables », le constat que les chiffres du passé récent sont tout aussi construits que les textes plus anciens, ont ouvert la voie à une quantification qui ne dédaigne pas les périodes longtemps qualifiées de « pré-statistiques », mais qui les traite avec ni plus ni moins de prudence que le contemporain.

Décrire et compter, compter et écrire

Deuxième raison de venir à la quantification, à côté du simple volume des sources : la volonté d'ajouter de nouvelles cordes argumentatives à son arc, de trouver les moyens d'armer, d'une manière inhabituelle, une intuition pour l'éprouver, puis pour convaincre ses collègues. La quantification peut, tout simplement, fournir des réponses raisonnablement

sûres à nos questions, qu'elles soient d'ordre descriptif ou explicatif. Comme l'espérait Jean-Yves Grenier, quantifier devient l'occasion d'explicitier, donc d'affiner une hypothèse. C'est ainsi, par exemple, que la proposition « la famille pèse sur l'institution » doit, pour être testée dans un cas précis, être décomposée en deux autres au moins, inégalement justes : « être de la famille aide à entrer dans l'institution » et « des familles y forment à chaque instant des minorités agissantes¹⁰ ». La spécification de la problématique, indispensable pour avancer sur le chemin de la quantification, importe ici presque autant que le résultat, fourni par une analyse de réseaux. De simples comptages permettent aussi de montrer que les peintres d'avant-garde exposés à Paris au début du XX^e siècle ne sont pas particulièrement étrangers ou que les associations de consommateurs de la même époque ne sont pas particulièrement féminines¹¹. Il reste alors à se demander comment est née la représentation contraire chez bien des contemporains – ce qui peut, du reste, mener à d'autres quantifications, qui pondèrent, par exemple, les pratiques par leur visibilité.

Tout cela conduit à une évolution que n'avait pas anticipée Grenier : une intégration croissante entre qualitatif et quantitatif, qui n'a rien d'évident ni de facile¹². Il est temps, en effet, de mettre fin à la dichotomie quantitatif *versus* qualitatif, sans s'en tenir au slogan de dépassement des oppositions. En effet, l'articulation entre qualitatif et quantitatif prend trop souvent, en histoire comme ailleurs, la forme d'une juxtaposition, conservant des chapitres dédiés, des formes d'écriture différentes, un rôle de simple cadrage ou de contextualisation pour le quantitatif ou, au contraire, de pure illustration pour le qualitatif. Pourtant, il y a bien des manières de les faire dialoguer pour produire du nouveau, qu'il s'agisse de mener une critique poussée des sources avant de quantifier et pour mieux le faire, ou bien de rebondir

d'un résultat quantitatif simple vers une nouvelle question au premier abord moins facile à formaliser.

Marier quantification et travail dit qualitatif, c'est donc surtout mélanger les résultats obtenus par les deux voies, les faire se répondre les uns aux autres, se confirmer ou dissoner de manière féconde. L'articulation des méthodes se fait dans l'écriture qui, sans suivre nécessairement le strict chemin de l'alternance, peut à la fois expliciter en amont la construction d'un tableau par la présentation des sources qui le composent, puis en faire comprendre le sens en aval par l'exposé d'un cas, ou le nuancer d'ailleurs. Prenons l'exemple des 991 Juifs de Lens étudiés pendant la Seconde Guerre mondiale. Parmi eux, 660 quittent le bassin lensois entre 1940 et l'automne 1942 ; or plus la date de départ est tardive, moins le risque d'arrestation est important. La question n'est pas tant, en effet, la date du départ que les destinations choisies pour partir et, notamment, le moment du basculement dans la clandestinité. La plupart des partants de la dernière heure (dans l'année 1942) quittent Lens pour se cacher et changent, dès lors, d'identité. *A contrario*, le caractère relativement officiel de leur départ rend ceux qui fuient dès 1940 plus vulnérables. Le cas d'Oscar Wajblum, qui écrit au maréchal Pétain le 29 juillet 1942 pour tenter d'obtenir l'autorisation de passer la ligne de démarcation, vient rappeler que la question d'un « bon moment » pour partir n'a pas la même signification pour l'historien d'aujourd'hui et le Juif lensois d'alors, qui ne sait pas, par définition, la « fin de l'histoire¹³ ».

Loin de plomber nécessairement l'écriture historique ou d'imposer des formes propres d'autorité scientifique à base de tableaux ou de courbes, la quantification peut ainsi s'intégrer sagement dans l'argumentation et la narration, comme un moment parmi d'autres – mais un moment qui peut être décisif dans leur avancée.

Nouveaux usages, nouveaux outils

Dernier constat sur les formes nouvelles de quantification : un glissement de ce que recouvre le terme, de la mesure et du chiffre vers la preuve et la forme. Charles Tilly, dans un texte précieux inventoriant de nombreux travaux originaux, soulignait ainsi que le terme de « formel » ou de « formalisation », plus large, serait plus approprié que celui de « quantitatif » ou de « quantification » pour désigner un ensemble de techniques qui donnent lieu à la production de représentations très variées du social, n'impliquant pas toujours des chiffres : cartes, schémas, graphes, etc¹⁴.

Ce que ces techniques ont en commun, c'est l'explicitation d'un certain nombre de procédures : celles qui visent à abstraire des données à partir de sources, puis celles qui visent à tirer de ces données des représentations plus abstraites encore de processus historiques – et souvent à utiliser ces représentations pour appuyer ou invalider des hypothèses. L'histoire labroussienne, toute quantifiée qu'elle était, n'était pas toujours très explicite sur les étapes de construction des données et des résultats¹⁵. Son objectif restait de savoir « combien », avec la meilleure précision possible : quel prix, combien de galériens, combien d'habitants en France en 1789, d'où une ambition d'exhaustivité visant non pas à rendre compte d'une diversité de trajectoires, mais bien à n'oublier personne dans la « pesée globale¹⁶ ».

Cette métaphore de la pesée, cette ambition d'exhaustivité, cette association entre formalisation, chiffre (de préférence avec décimales) et mesure exacte, les historien(ne)s d'aujourd'hui qui pratiquent la formalisation ont dû les désapprendre ou ont choisi de les ignorer au profit d'autres usages, qui ont plutôt à voir avec l'abstraction. Il faut en effet souvent cacher, provisoirement, une partie de la réalité reconstituée pour mieux y discerner des motifs qui font

sens. Les techniques de *datamining*, qui vise à trouver des ordres locaux dans des données complexes, mettent en évidence ces motifs, voire de véritables structures. Il faut aussi abstraire pour définir des hypothèses sur des processus historiques qui soient assez précises pour qu'elles puissent être testées – le test n'étant en général pas la fin du travail historique, mais ouvrant sur de nouvelles questions.

La place différente de la formalisation dans le déroulement d'une recherche historique implique le recours à des outils différents de ceux qui avaient été routinisés dans la période d'enthousiasme pour le quantitatif. Ainsi, il s'agit très rarement des séries temporelles recensées par Grenier, celles qui aboutissaient à une page dépliant des *Annales*, sur papier sémi-logarithmique, avec une courbe de tendance et des fluctuations conjoncturelles. D'autres routines ont la vie plus dure, mais finissent par laisser place à des innovations. La prosopographie, d'ailleurs souvent peu quantifiée¹⁷, est une de ces routines. Mais l'ouvrage précité sur les Juifs de Lens montre l'intérêt d'adapter la démarche prosopographique à d'autres groupes que les membres d'une élite institutionnelle. De plus en plus de travaux prennent aussi le parti de la comparaison, entre le groupe étudié et d'autres ou au sein même de ce groupe, ainsi que du changement d'échelle et des allers et retours, dans l'écriture, entre quantification et récits de trajectoires individuelles, ce qui évite la platitude des énumérations de pourcentages. Ainsi, Séverine Sofio met à l'épreuve des faits les discours, peu nombreux et dévalorisants, parlant des plasticiennes dans la littérature du XIX^e siècle, en menant l'analyse comparée des carrières des hommes et des femmes exposant au Salon entre 1791 et 1848. Son étude qui confronte descriptions contemporaines et pratiques porte sur un corpus massif (plus de 1 000 artistes) tout en ayant recours à des plongées individuelles, *via* l'étude de correspondances, ce qui lui permet

de montrer que le genre n'est finalement pas une catégorie si pertinente pour décrire le monde des arts plastiques au début du XIX^e siècle¹⁸. De même, un certain usage, très statique, mécaniquement bourdieusien ou peu attentif aux individus, de l'analyse factorielle¹⁹ semble se raréfier, sans que la méthode même soit disqualifiée. Étudiant les Algériens qui sollicitent un regroupement familial dans l'après-guerre, Muriel Cohen lui accorde un rôle d'auxiliaire, utile pour construire des typologies, avec l'assistance d'une classification automatique, qui incite aussi aux allers et retours avec les cas individuels en désignant les « parangons », individus typiques de chaque classe²⁰. Elle met ainsi au jour des trajectoires très standardisées et d'autres plus exceptionnelles, comme celles de migrants arrivés jeunes en France et relativement qualifiés.

Au-delà de ces renouvellements des pratiques anciennes de quantification, on observe surtout une diversification tant des objets quantifiés, des questions formalisées et des sources mobilisées que des techniques utilisées pour produire des résultats. Seul le passage par une base de données construite au plus près de la source est un dénominateur commun. C'est bien souvent à ce moment précis de la recherche, celui de la saisie – qui n'implique, en soi, aucun logiciel spécifique –, que les savoir-faire, plutôt historiens que « quantitativistes », sont mis à l'épreuve²¹. On ne saurait ici trop insister sur l'importance de cette étape, généralement laissée de côté par les cours et les manuels, alors qu'elle est cruciale. C'est elle qui assure que la construction des données suit les pratiques de critique des sources qui fondent la discipline qu'est l'histoire. Et en pratique, c'est aussi lorsque l'on saisit soi-même ses données que naissent les intuitions qui permettront à la fois de mieux les structurer, de leur poser des questions plus neuves et plus adaptées et, *in fine*, d'écrire de la bonne histoire – ce qui pose la question des formes de

sous-traitance imposées par les très grands corpus, ou par les modes actuels de financement de la recherche, et de la manière de conserver, si on doit s'y résoudre, ce contact avec la source. Ainsi, c'est en cherchant à repérer les rescapés des camps parmi les centaines de dossiers des personnes aidées dans l'immédiat après-guerre par le Casip-Cojasor, une fondation caritative juive créée en 1809, que Laure Fourtage s'aperçoit que les « déportés », sur lesquels devait au départ porter sa recherche, ne constituent pas une catégorie identifiable dans la source, ni pertinente pour l'institution²².

L'intervention d'un jugement proprement historien lors des étapes de saisie, puis de codage des données – et le simple fait de bien distinguer ces deux étapes, permis par les progrès de l'informatique – est sans doute le principal dénominateur commun des nouveaux usages de la quantification. Il implique que celle-ci s'inscrit pleinement dans le tournant constructiviste pris par les sciences sociales. Loin d'imposer l'usage de catégories exclusives, simplettes ou anachroniques, un objectif de quantification peut, en effet, puissamment inciter à réfléchir sur la substance et le sens de nos catégories ou de celles des sources. Paul-André Rosental en a donné un exemple pionnier en prenant au sérieux une des questions que posent le plus souvent les sources historiques, celle du codage des noms de métiers. Refusant de n'en faire qu'une opération technique non discutée, il se livre à un jeu de recodages : les résultats diffèrent suivant qu'on permet ou non une seule mention par individu, suivant qu'on utilise des catégories socioprofessionnelles pré-existantes ou qu'on infère des proximités entre métiers à partir de présences croisées comme témoins au mariage. Ce jeu n'est pas relativiste : de ces différentes facettes émerge une image plus subtile, mais pas moins nette, de la stratification sociale étudiée²³. Ainsi, pour répondre aux critiques élevées à l'encontre d'une histoire sociale reprenant trop

aveuglément les catégories de sociologues ou bien celles, administratives, de ses sources, deux voies complémentaires et non concurrentes s'ouvrent aujourd'hui : l'interrogation sur la construction des sources peut prendre la forme d'une histoire de l'administration, mais elle peut également s'incarner dans un retour informé à l'histoire sociale, y compris quantitative, fondée sur l'expérimentation de différents codages possibles et leurs effets heuristiques.

En amont et en aval de ces étapes cruciales que sont la saisie et le codage de l'information historique, la diversification des sources et des techniques de traitement n'a pas suivi les voies envisagées par Grenier, qui pensait surtout à l'histoire économique et évoquait des manières de modéliser, donc d'abstraire radicalement des raisonnements, qu'elles passent par un recours aux concepts de chaos ou de fractales ou par la théorie des jeux. Plus exactement, si des travaux dans cette veine existent, ils ont rarement fait école²⁴. Ce qui a fleuri, ce sont des traitements pas toujours plus simples à apprendre – ainsi, Anne-Sophie Bruno utilise des régressions multi-niveaux pour rendre compte des carrières professionnelles d'originaires de Tunisie dans la France des Trente Glorieuses²⁵ –, mais plus directement orientés vers l'analyse de données empiriques, qu'il s'agisse de simple comptage, de *datamining*, c'est-à-dire de recherche de structures dans des données complexes, ou de test d'hypothèses. C'est, par exemple, l'*event history analysis* qui a permis à Jérôme Krop de discuter le poids réel, sur les promotions d'instituteurs parisiens de la Belle Époque, de recommandations personnelles qu'il retrouvait régulièrement dans ses sources ; c'est l'analyse de réseaux qui a fait découvrir à Isabelle Rosé, pourtant déjà auteure d'une biographie d'Odon de Cluny, de nouveaux aspects de l'évolution de ses entourages au cours d'une longue vie riche en voyages²⁶. Il semble enfin aujourd'hui possible d'intégrer le continuum

temporel au sein d'analyses quantifiées, sans l'isoler comme une simple « variable ». L'œuvre d'Andrew Abbott a ainsi permis le renouveau de la sociologie historique en suscitant une prise en compte du temps non restreinte au schème de la genèse, c'est-à-dire de la recherche des origines de pratiques ou d'institutions actuelles – un schème qui reste encore dominant en France, notamment dans le cadre des études dites « socio-historiques ». En remettant au contraire l'histoire à l'endroit, en envisageant des trajectoires professionnelles et biographiques dans leur totalité et à partir du début, il propose des perspectives d'analyses quantifiées particulièrement stimulantes pour lire, interpréter et utiliser les données prosopographiques²⁷.

Au-delà de cette énumération, l'important est que des historien(ne)s en soient venu(e)s à utiliser telle ou telle méthode non pas parce qu'ils-elles la considéraient comme la meilleure pour l'histoire en général, ou parce que c'était la seule qu'ils-elles avaient apprise, mais parce qu'elle s'avérait la mieux adaptée à leurs sources et à leurs questions : loin de se sentir marié(e)s, à vie, à une méthode ou à une autre, ils-elles en ont même souvent marié plusieurs sur un même objet²⁸. Ce pluralisme méthodologique, permis par la multiplication des manuels et tutoriels non mathématisés, y compris en français²⁹, demande certes des efforts de formation ; mais ils sont récompensés lorsqu'ils permettent d'exploiter à plein la richesse de sources historiques péniblement lues et transcrites.

Claire Lemerrier et Claire Zalc