



La statistique dans la cité - décembre 2016

Lettre bimestrielle du groupe « Statistique et enjeux publics »

Sommaire du n°2 – décembre 2016

Editorial : *Vœux*

Droit et statistique : *Vers un nouveau règlement statistique européen*

Méthodes : *A propos de l'élection de Donald Trump*

Outils : *Le répertoire statistique des logements : un outil prometteur pour la démographie*

Travaux : *Perception vs mesure*

Agenda : *Les Cafés de la statistique ont repris*

Humeur : *Difficile cohérence...*

Éditorial

L'année 2017 est à nos portes. Elle sera riche d'événements politiques avec l'élection présidentielle et les législatives qui suivront. Les sondages d'opinion, les sondages de popularité des candidats et les sondages préélectoraux occuperont beaucoup les médias, avec leur cortège de commentaires et, bien sûr, de mises en question des méthodes.

Le groupe « Statistique et enjeux publics » suivra avec attention les débats techniques entourant les sondages. Il prévoit d'organiser en mai un Café de la statistique sur les sondages politiques afin de rappeler leurs vertus et leurs limites. Ce Café prolongera une longue série de soirées puisque le Café n° 100 (déjà !) sera célébré le 27 mars.

Se voulant informative, *La statistique dans la cité* sera attentive à tout ce que la statistique peut apporter d'éclairage au fonctionnement de notre société si complexe, mais aussi à tous les débats qu'elle peut soulever par ses méthodes et ses productions. Elle compte aussi sur ses lecteurs pour lui signaler les faits méritant à leurs yeux d'être relevés.

L'ambition de cette lettre bimestrielle, au-delà de l'information, est aussi de susciter des collaborations des professionnels de la statistique pour des actions communes en faveur de l'utilité publique de leur discipline. La rédaction lira par ailleurs avec intérêt les critiques et suggestions de tous. A toutes ses lectrices, à tous ses lecteurs, *La statistique dans la cité* souhaite une excellente année 2017.

Pour nous écrire : sep@sfds.asso.fr

Droit et statistique

Vers un nouveau règlement statistique européen

Le Comité du système statistique européen (Comité SSE) propose un règlement instituant un cadre commun en ce qui concerne les données individuelles recueillies par échantillon pour les statistiques européennes relatives aux personnes et aux ménages (IESS, pour *Integrated European Social Statistics*). Il en a adopté le texte le 24 août dernier.

Cette proposition vise à établir un cadre et une périodicité unifiés qui permettront de rationaliser des processus auparavant définis par des règlements différents, générant des charges de travail erratiques et peu prévisibles d'une année sur l'autre.

Le projet de règlement a pour champ les collectes concernant les marchés du travail des pays de l'Union européenne, les revenus et les conditions de vie, la santé, l'éducation, la formation et l'usage des technologies de l'information et de la communication.

Le projet doit être soumis au Conseil et au Parlement européens pour discussion et approbation. Les règlements d'application qui suivront devraient permettre de mieux harmoniser les concepts utilisés en vue d'améliorer la comparabilité des données entre pays.

Méthodes

A propos de l'élection de Donald Trump :

Les instituts de sondage américains ont-ils vraiment failli ? Creusons un peu

De fait, au niveau du nombre total de voix obtenues au niveau fédéral, les sondages sont loin d'avoir failli puisqu'il semble bien, d'après les derniers comptages, qu'Hillary Clinton ait obtenu plus de deux millions de voix de plus que son concurrent élu.

Mais bien entendu, compte tenu du système électoral américain, tout se joue au niveau des États, et surtout au niveau des quelques sept ou huit États clés (les fameux « Swing States »). Des sondages sont aussi bien sûr effectués dans les États et en particulier dans les « Swing States ». Or dans chacun de ces derniers, les sondages effectués les jours précédant le 8 novembre donnaient une chance sur deux de gagner à chacun des deux principaux candidats. Sans doute le raisonnement fait par tous les commentateurs a-t-il été le suivant : la probabilité que l'un ou l'autre des deux principaux candidats gagne chacun de ces États était de 50 % ; mais la probabilité qu'un des deux candidats gagne simultanément dans les huit États considérés était très faible (0,5 à la puissance 8) et donc, puisqu'il suffisait qu'Hillary Clinton gagne dans seulement deux ou trois de ces « Swing States » pour obtenir la majorité des grands électeurs au niveau fédéral, les commentateurs ont pu estimer que sa probabilité de gagner l'élection au niveau fédéral était sensiblement supérieure à 0,5. Mais ce raisonnement aurait supposé pour être exact que les événements dans les huit urnes que constituent les huit « Swing States », aient été indépendants. Ce qui n'était évidemment pas le cas, et la probabilité que l'un ou l'autre des deux candidats gagne simultanément tous ces États était donc en fait plus voisine de 0,5 que de zéro. Et donc les commentateurs auraient dû être prudents. Ce sont eux, plutôt que les instituts de sondage américains, qui ont failli.

De façon générale, des résultats différents de ceux que pourraient laisser apparaître les sondages peuvent se produire, avec tout système électoral, quand les intentions de vote sont proches les unes des autres, et ce ne sont pas les sondages qu'il faut incriminer mais l'utilisation qui en est faite, en particulier dans le cas américain.

Source : Résultats des sondages effectués au niveau de chaque État

https://fr.m.wikipedia.org/wiki/Liste_de_sondages_sur_l'élection_présidentielle_américaine_de_2016

Outils

Le répertoire statistique des logements : un outil prometteur pour la démographie

L'utilisation des données administratives pour produire des statistiques est recommandée par les organisations internationales et elle est de pratique courante dans les pays nordiques. Et aux Pays-Bas, par exemple, les recensements de population traditionnels sont depuis plusieurs décennies remplacés par des appariements de fichiers administratifs.

En France, la statistique d'entreprise recourt depuis longtemps aux fichiers fiscaux. Mais en matière de statistique démographique et sociale, l'interrogation directe des ménages reste le moyen le plus couramment employé.

Le répertoire statistique des logements (RSL), développé par l'Insee depuis 2011, s'appuie au contraire sur des fichiers d'origine fiscale : fichier de la taxe d'habitation et fichiers d'imposition des personnes. Ces données, mises sous une forme de fichiers statistiques exhaustifs et anonymisés, font l'objet d'une géo-localisation très précise. Il en résulte un outil de connaissance finement localisée du bâti, des logements, des personnes et de leur mobilité résidentielle.

De premières études prometteuses issues du RSL sont parues en 2015 ; la cohérence avec les autres outils de la statistique sociale est très bonne. Une fois le projet achevé, il est prévu que cette source soit disponible pour les chercheurs dans un centre d'accès sécurisé.

Références :

- Une étude publiée utilisant le RSL : Insee-Analyses Pays de la Loire n° 34 Juin 2016 « Loire-Atlantique, Des motifs de déménagement différents selon les territoires et les distances ».
- Discussion de la nouvelle source au Conseil national de l'information statistique (Cnis) : compte rendu de la réunion de la commission Territoires du 10 juin 2015.

Travaux

Perception vs mesure

Les Français passent pour chroniquement pessimistes. Abondamment commenté déjà, ce thème de la sinistrose est repris par un récent rapport de France Stratégie (ex Commissariat général du Plan) ; <http://www.strategie.gouv.fr/sites/strategie.gouv.fr/files/atoms/files/rapport-lignes-de-faille-ok.pdf>

Ce rapport met en regard ce que les Français pensent de la situation présente et ce que les enquêtes statistiques en disent. Confirmant qu'ils voient en noir ce qui n'est que gris. Les perceptions individuelles s'écartent souvent de la mesure statistique :

- chacun est sensible à ce qui lui est proche ou familier et méconnaît souvent ce qui est loin mais participe pourtant à la mesure d'ensemble ;

- les concepts pertinents au niveau collectif ne le sont pas toujours au niveau individuel, d'où des malentendus. Par exemple, il n'est pas sûr qu'une élévation générale des qualifications diminue le taux de chômage, mais il est probable que j'accroîtrai mes chances d'emploi en élevant ma propre qualification.

Cet écart, classique et pourtant souvent méconnu, avait déjà fait l'objet de deux Cafés de la statistique : "Observation des prix et perception du consommateur" (2006) et "Statistique et sentiment personnel" (2010). Ici, France Stratégie n'examine pas l'écart en question sous l'angle méthodologique, celui d'une "erreur" de perception, mais comme confirmation et mesure du fait de psychologie collective que constitue ladite sinistrose. Autrement dit, une troisième cause d'écart : l'autonomie du jugement envers le fait observable. Son rapport a fourni l'occasion d'un nouveau Café de la Statistique, le 8 novembre dernier, dont on peut trouver la vidéo et le compte rendu à l'adresse suivante :

<http://www.sfds.asso.fr/321-Les-Cafes-de-la-Statistique>

Globalement, il ressort du rapport que la confiance des Français envers les institutions fait défaut et que de nouvelles règles de vie en commun sont à imaginer.

Agenda

La saison 2016-2017 des Cafés de la statistique a commencé le 8 novembre avec un débat sur les lignes de faille qui traversent la société française. Le 6 décembre, c'est le développement de la personne dans la petite enfance qui était en question. Les Cafés suivants auront pour thème :

- le mardi 10 janvier, le développement économique de l'Afrique ; avec **Serge Michailof**, Iris – Sciences Po ;
- le mercredi 22 février, l'évolution des emplois en relation avec les innovations technologiques et la robotique ; avec **Dominique Méda**, Paris-Dauphine ;
- le lundi 27 mars (Café n° 100 !), l'exploitation des données massives ; avec **Valérie Peugeot**, sociologue, chercheuse à Orange Labs et commissaire à la Cnil.

Humeur

Difficile cohérence...

Dans le supplément *Eco&entreprise* du *Monde* daté du 23-24 octobre, page 1, on lit que le projet de budget de la Russie pour 2017 avoisine 1 600 milliards (Mds) de roubles (235 Mds d'euros) ; que les fonds secrets ont augmenté de 800 Mds de roubles selon le correctif apporté à l'exercice 2016 ; que les dépenses militaires réelles dépassent 4 200 Mds de roubles (5,3 % du PIB) et que la somme globale des fonds secrets russes est estimée à 3 700 Mds de roubles représentant plus de 20 % du budget.

Tout cela laisse perplexe :

1 – Au taux de change actuel, 1 600 Mds de roubles représentent 23,5 Mds € environ (et non 235) ;

2 – Les dépenses réelles pour 2016 étant estimées à 16 100 Mds de roubles, on est en droit de penser que le budget 2017 est de 16 000 Mds (et non 1 600) ;

3 – 20 % du budget représenterait donc 3 200 Mds de roubles. On approche des 3 700 Mds annoncés pour les fonds secrets ;

4 – Le PIB de la Russie est d'environ 74 250 Mds de roubles et 5,3 % correspondent à 3 935 Mds, montant pas trop éloigné des 4 200 Mds de dépenses militaires réelles.

Ouf ! Le lecteur arrive à une certaine cohérence, mais au prix de recherches d'informations extérieures à l'article ! A quand un logiciel de correction automatique des incohérences de chiffres dans les rédactions ?

Responsable de l'infolettre : Marion Selz, présidente du groupe SEP

Rédacteur en chef : Alain Godinot

Secrétaires de rédaction : Jean-Louis Bodin, Jean-Pierre Le Gléau

Webmestre : Érik Zolotoukhine

Ceci est un message automatique, merci de ne pas y répondre.

Si vous ne souhaitez plus recevoir d'informations de la SFdS, veuillez vous [connecter à votre compte](#) et modifier vos paramètres.