

Accès aux données de la statistique publique et ressources de données en ligne

Alina Danciu, Alexandre Mairot

▶ **To cite this version:**

Alina Danciu, Alexandre Mairot. Accès aux données de la statistique publique et ressources de données en ligne. Quantilille, École d'été de Lille en Méthode Quantitative des Sciences Sociales, Jun 2018, Institut d'études politiques de Lille, France. hal-03570939

HAL Id: hal-03570939

https://sciencespo.hal.science/hal-03570939

Submitted on 15 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accès aux données de la statistique publique et ressources de données en ligne

École d'été de Lille

en Méthodes Quantitatives des Sciences Sociales 2018

Alina DANCIU, Alexandre MAIROT

(Centre de données socio-politiques, Sciences Po)

Les bases de données internationales

- Eurostat : Office statistique de l’Union européenne 

Eurostat propose toute des données que les pouvoirs publics, les entreprises, le secteur de l'éducation, les journalistes et les citoyens peuvent utiliser.

- OCDE : Organisation de coopération et de développement économiques 

La mission de l’OCDE est de promouvoir les politiques qui amélioreront le bien-être économique et social partout dans le monde.

- ISSP, International Social Survey Program 

L'enquête ISSP est réalisée dans d’environ 50 pays. Le thème de cette enquête change chaque année et revient périodiquement pour permettre la comparaison.

IPUMS International, Integrated Public Use Microdata Series 

L’IPUMS est dédié à la collecte et à la distribution des données du recensement provenant de 89 pays.

Les enquêtes européennes

- SHARE, Survey of Health, Ageing and Retirement in Europe

L'enquête est réalisée tous les deux ans depuis 2004 et se déroule désormais dans 27 pays européens. Les données portent sur l'état de santé des répondants et de leurs proches ainsi que sur leur situation sociale (famille, entraide, réseaux sociaux) et économique (emploi, retraite, patrimoine).

- ESS, European Social Survey

L'European Social Survey est destinée à mesurer les comportements et les attitudes des citoyens de pays membres ou non de l'Union européenne sur des thèmes socio-politiques. 22 pays ont participé à la 1re édition (28 pour la 6e édition).

- GGP, Generations and Gender Program

Le Generations and Gender Program fournit des données sur les familles et sur les trajectoires de vie des individus. La 1re édition date de 2004. À ce jour, l'enquête a été menée dans 20 pays européens et 4 autres États.

Catalogues de statistiques européennes

- CIMES, Centralising and Integrating Metadata from European Statistics

Le CIMES fournit un catalogue des données diffusées à des fins de recherche par les instituts nationaux de statistiques à travers l'Europe.

- MISSY, Microdata Information System

MISSY fournit un catalogue des données diffusées par les instituts nationaux de statistiques à travers l'Europe portant en particulier sur l'éducation, l'innovation, le travail, les revenus, les conditions de vie…

La statistique publique française

- Le recensement de la population 
  - Évolution de la population
  - Couples, Familles, Ménages
  - Logements
  - Diplômes, formation
  - Population active, emploi, chômage
  - Caractéristique de l'emploi
  - Étrangers, immigrés
  - Série historique depuis 1968
- Des données infracommunales 
  - Activités
  - Diplômes, formation
  - Couples, familles, ménages
  - Logements
  - Population
- Une banque d'indicateurs statistiques 
- Des cartes interactives 

Les enquêtes issus de la recherche française

- Les grandes enquêtes sociodémographiques produites et diffusées par l'INED 

Thèmes : couples, famille, sexualité, naissance, fécondité, discrimination, migration, genre, populations du monde, Les données socio-politiques en particulier les enquêtes électorales

- Des données socio-politiques, en particulier les enquêtes électorales du CEVIPOF par le CDSP 

Enquêtes de recherche sur des thématiques diverses (enquêtes des écoles d'ingénieurs, enquêtes « Presse Gaie », enquête longitudinale sur la socialisation professionnelle des gardiens de la paix, enquêtes sur l'image de la science, enquêtes du panel ELIPSS

Enquêtes de la statistique publique et issues de la recherche française

Quetelet PROGEDO Diffusion

- Archives de Données Issues de Statistiques Publique (ADISP)
- Centre de données socio-politiques (CDSP)
- Institut national d'études démographiques (INED)
- Centre d'accès sécurisé aux données (CASD)

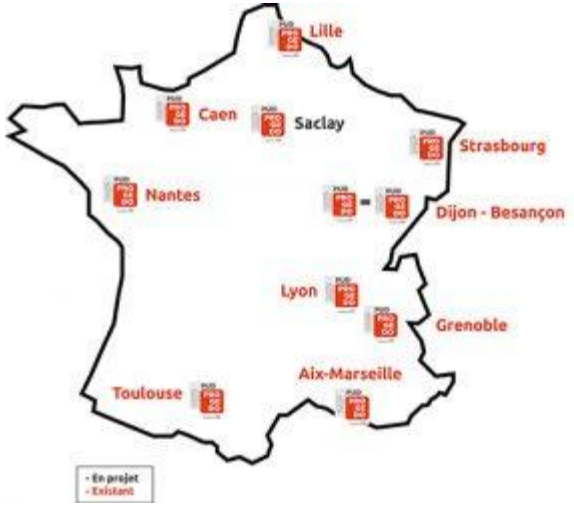
Plates-formes universitaires de données

Les plates-formes universitaires de données (PUD) visent à stimuler l'utilisation des données mises à disposition par PROGEDO. Centrées sur le développement des compétences, les PUD prolongent l'action de PROGEDO au niveau régional et apportent un soutien direct aux utilisateurs de données d'enquêtes par :

– l'aide à la recherche, à l'acquisition et à la mise en œuvre des fichiers de données,

– la formation à la compréhension, à l'utilisation et au traitement de ces données,

– la formation à l'utilisation d'outils informatiques et statistiques.



La recherche de données

- Outil de recherche de données en sciences sociales 

Il permet une recherche d'enquêtes au sein d'une base documentaire en langue française et oriente l'utilisateur vers l'unité de diffusion correspondante.

- Base de questions 

La base de questions permet la recherche dans les textes des questions (question + modalités) et des variables. Elle comporte 911 enquêtes.

- Catalogues de l'ADISP , de l'INED  et du CDSP 

Les serveurs Nesstar de l'ADISP, de l'INED et du CDSP permettent un accès en ligne, direct et gratuit aux enquêtes diffusées.

Exemple 1: Enquête sur les pratiques culturelles des Français

- ORDS, « Pratiques culturelles »



- Base de questions

- Catalogue du CDSP

Exemple 2: Enquête sur les formations des choix électoraux

- ORDS, « choix électoraux »



- Base de questions

- Catalogue du CDSP

La commande de données

- Création d'un compte et d'indentification 
- Demande de données :
  - Description du projet de recherche
  - Signature et envoi d'un formulaire d'engagement

ADISP :
Les originaux du formulaire d'engagement et de la licence de droit d'usage (pas de copie) par voie postale uniquement.

CDSP :
Le formulaire d'engagement à renvoyer par voie postale, électronique ou fax.

INED :
Le formulaire original d'engagement (pas de copie) à renvoyer
par voie postale uniquement.

- Téléchargement des données (après acceptation du diffuseur)

N’oublions pas la citation

Les diffuseurs de données fournissent un modèle de citation

Sans citation de votre part, il nous est impossible de savoir si les données sont ré-utilisées

Difficile à convaincre les chercheurs et institutions à partager leurs données

Divers travaux montrent que les articles qui citent les données ré-utilisées ont plus de chances d’être citées à leur tour

Exemple : Henneken, Accomazzi, ‘Linking to Data - Effect on Citation Rates in Astronomy’ http://arxiv.org/pdf/1111.3618v1.pdf

Dépôt de données et leur réutilisation

Au sein de Quetelet PROGEDO Diffusion, l'ADISP et le CDSP accompagnent les institutions et chercheurs qui souhaitent mettre à disposition leurs données et enquêtes en vue de leur réutilisation par la communauté scientifique et académique.

