

Quantification et données en SHS ou comment sensibiliser chercheurs et étudiants à une utilisation éclairée des données quantitatives ?

Frédérique CORNUAU, MCF en démographie, CLERSE-Université de Lille

Marie CROS, IE, PUDL, Université de Lille

Des données...

Différents types de données

- Enquêtes / données administratives
- Statistique officielle / organismes privés / travaux de chercheurs
- Données agrégées / données détaillées

Multiplicité...

- Sources
- Producteurs
- Thématiques
- Espaces - échelles
- Temporalités



Plateforme Universitaire de Données de Lille

“ Favoriser et développer l'utilisation des données quantitatives en SHS ”

- Trouver des données
- Former aux logiciels
- Aider à analyser



Sensibiliser à une utilisation éclairée des données quantitatives

- 1) Quelles données pour quel questionnement ?
- 2) Règles déontologiques et éthiques
- 3) Règles méthodologiques et statistiques



1. Quelles données pour quel questionnaire ?



- 1) Orienter vers la source pertinente
- 2) Nécessité d'homogénéiser
- 3) Importance des métadonnées
- 4) Données et outils adaptés aux compétences des utilisateurs

Etudier l'évolution socio-démographique des quartiers

Caractériser un territoire selon des variables disponibles à des échelons différents : vote FN, degré ruralité, taux de chômage...

- Formation, qualification professionnelle - 2003
- Métadonnées
 - ⊕ Description de l'étude
 - ⊕ Description des fichiers de données
 - ⊕ Autre documentation
- Description des variables
 - ⊕ Tronc commun
 - ⊖ Questionnaire individuel
 - ⊕ Etat civil et situation familiale
 - ⊕ Situation professionnelle au dernier emploi (s'il est postérieur à 1998)
 - ⊕ Situation professionnelle en 1998
 - ⊕ Situation professionnelle lors

Jeu de données: Formation, qualification professionnelle - 2003

Variable FI: Occupation professionnelle (principale) actuelle

| Valeurs | Catégories | N | |
|---------|--|-------|-------|
| 1 | Exerce une profession | 25784 | 65,6% |
| 2 | Chômeur (inscrit ou non à l'ANPE) | 2966 | 7,5% |
| 3 | Étudiant, élève, en formation, en stage non rémunéré | 2825 | 7,2% |
| 5 | Retraité (ancien salarié) ou pré-retraité | 3092 | 7,9% |
| 6 | Retiré des affaires (ancien agriculteur, ancien artisan, ancien commerçant) | 325 | 0,8% |
| 7 | Personne au foyer (y compris congé parental) | 3075 | 7,8% |
| 8 | Autre inactif (y compris les personnes ne touchant qu'une pension de reversion et les personnes invalides) | 1218 | 3,1% |

STATISTIQUES SOMMAIRES

| | |
|-----------------------------------|-------|
| Cas valides | 39285 |
| Cas manquants | 0 |
| Cette variable est alphanumérique | |

2. Déontologie et éthique



- 1) Respect du secret statistique
- 2) Données personnelles & CNIL
- 3) Anonymat et confidentialité

« Constitue une *donnée à caractère personnel* toute information relative à une personne physique identifiée ou qui peut être identifiée, directement ou indirectement(...) »

« Les *données sensibles* sont celles qui font apparaître, directement ou indirectement, les origines raciales ou ethniques, les opinions politiques, philosophiques ou religieuses ou l'appartenance syndicale des personnes, ou sont relatives à la santé ou à la vie sexuelle de celles-ci. »

Art. 2 de la loi "Informatique et libertés"

Etudier la nationalité détaillée des habitants d'une commune.

Collecter des données personnelles par questionnaire

Etudier une population de taille restreinte

3. Règles méthodologiques et statistiques : pour une fiabilité des résultats



- 1) Interpréter des évolutions
indicateurs / zonages / collecte / formulation
- 2) Comparabilité
champ couvert / contexte
- 3) Représentativité
- 4) Significativité

*Etudier l'évolution
du statut dans
l'emploi*

*Etudier l'évolution
du taux d'activité à
travers les enquêtes
emploi*

*Comparer l'emploi
dans les pays
européens*

Conclusion

- 1) Aide personnalisée
- 2) Actions globales vers l'enseignement
- 3) Actions globales envers la recherche

Sources & crédits

- *Guide du secret statistique, INSEE, 2010*
- *www.cil.cnrs.fr*
- *Kych.A, Le recensement rénové de la population. - Tabulations et secret statistique, que faire ? 2011*
- *Régnier-Loilier.A, Conditions de passation et biais occasionnés par la présence d'un tiers sur les réponses obtenues à l'enquête Érfi, Economie et Statistique, 2007*
- *Crédit : SlidesCarnival icons*

Merci de votre attention !

Des questions ?

frederique.cornuau@univ-lille1.fr

marie.cros@univ-lille1.fr

<http://pudl.meshs.fr>