

# Guide de lecture de la cartographie des bases de données santé en France

# Objectif, démarche suivie et structure de l'excel de cartographie des données de santé en France

Zoom pages suivantes

L'objectif de ce fichier est de recenser les principales bases / jeux de données sur la santé en France

## Travaux effectués

- **Revue des principaux rapports** existants sur les données de santé en France
- Lecture des **sites internet des principaux gestionnaires** de données
- **Rencontre des principaux gestionnaires** de données de santé en France et de quelques agrégateurs

Fichier tableur  
(formats XLS, ODS et CSV)



## Structure du fichier

- Onglet « Introduction »
- Onglet « **Cartographie France** »
- Onglets « Annexes » :
  - Acronymes des gestionnaires
  - Critères d'ouverture

# Colonnes 1 et 2 : nous avons utilisé une typologie des données de santé appliquée à chacune des bases

## Utilisation d'une typologie en deux niveaux

### Typologie 1

Santé publique et épidémiologie

### Typologie 2

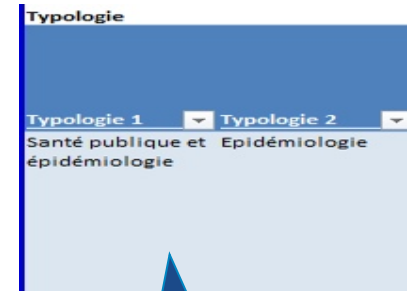
- Etat de santé
  - Epidémiologie
  - Environnement
  - Habitudes de vie
  - Inégalités de santé
  - Autres
- 
- Infrastructures
  - Services proposés
  - Personnel
  - Tarifs, honoraires
  - Infrastructure de recherche
  - Informations médicales
  - Autres
- 
- Activité des établissements de santé
  - Consultations
  - Biologie
  - Transports
  - Médicaments et dispositifs
  - Indemnités et accidents du travail
  - Activité des établissements médicaux-sociaux
  - Dépenses
- 
- Performance financière
  - Performance opérationnelle
  - Qualité

Offre de soins

Consommation de soins

Performance et opérations

## Colonnes 1 et 2 de la cartographie France



Chaque base de données est classée selon les deux niveaux de typologie

## Colonnes 3 à 5 : pour chacune des bases de données recensée, nous avons renseigné l'origine des données et leur(s) gestionnaire(s) actuel(s)

### Colonnes 3 à 5 de la cartographie France

#### Base, gestionnaire et origine des données

Bases/enquêtes/publications	Gestionnaire	Origine des données
Enquête Santé et Protection Sociale (ESPS)	IRDES	Echantillon de la population (panel de plus de 20 000 personnes sélectionnées grâce à l'EGB) : remontée des informations par questionnaire

#### Par base :

- Nom de la base
- Nom du gestionnaire principal

Origine des données et système de collecte<sup>1</sup>

<sup>1</sup> Cette colonne est parfois complétée en colonne 22 par des commentaires sur la couverture des données

# Colonnes 6 à 21 : nous avons étudié l'ouverture des données de chaque base aux niveaux agrégé et granulaire<sup>1</sup> selon 4 critères<sup>2</sup> d'ouverture

## Colonnes 6 à 13 de la cartographie France

DONNEES AGREGÉES

### Statut d'ouverture des données agrégées

Principales données	Liberté d'accès	Commentaires sur la liberté d'accès	Coût d'accès	Format de mise à disposition	Commentaire sur le format de mise à disposition	Facilité de réutilisation	Portail de mise à disposition
Tableaux de résultats sur la protection sociale des individus, les maladies et troubles de santé déclarés, les comportements, habitudes de vie et facteurs de risque, les recours et renoncements aux soins	Accès à tous	-	Gratuit	Non exploitable	Pdf	Explicité avec restriction	<a href="http://www.irdes.fr/Publications/Rapports2012/rap1886.pdf">http://www.irdes.fr/Publications/Rapports2012/rap1886.pdf</a>

Description des données mises à disposition

Revue des 4 critères d'ouverture<sup>2</sup> et commentaires:

- Liberté d'accès
- Gratuité d'accès
- Format exploitable
- Possibilité de réutilisation

Lien URL du portail de mise à disposition

## Colonnes 14 à 21 de la cartographie France

DONNEES GRANULAIRES

### Statut d'ouverture des données granulaires

Principales données	Liberté d'accès	Commentaire sur la liberté d'accès	Coût d'accès	Format de mise à disposition	Commentaire sur le format de mise à disposition	Facilité de réutilisation	Portail de mise à disposition
L'enquête santé et protection sociale (ESPS) de l'IRDES porte sur 6 thèmes principaux : les modes de protection sociale (obligatoires et facultatives), les maladies et les troubles de santé déclarés, les recours au médecin, la consommation de biens médicaux, les soins d'auxiliaires, l'hospitalisation, le renoncement aux soins et les opinions sur la santé	Accès restreint	Membres du conseil scientifique puis autres chercheurs	Gratuit	Exploitable	Txt, SAS	Explicité avec restriction	-

<sup>1</sup> Voir Annexe 1 ; <sup>2</sup> Détails en Annexe 2

# ANNEXES

# Nous avons distingué deux niveaux de granularité des données pour chaque base

Une même base de données peut fournir...

... des données granulaires

## Description des données

- **Données au niveau le plus fin qu'il est possible d'obtenir en fonction de l'origine de la donnée et du système de collecte**
- Par exemple :
  - Dans le cas d'une base contenant des données sur l'état de santé des patients, le niveau le plus fin est celui de l'individu, les données granulaires sont donc les données individuelles de chaque patient
  - Dans le cas d'une base contenant des données collectées sur les infrastructures de santé, le niveau le plus fin est celui de l'établissement, les données granulaires sont donc les données relatives à chaque structure

...et des données agrégées

- **Données obtenues en regroupant des données granulaires selon une ou plusieurs caractéristique(s) commune(s)**
- Par exemple :
  - Les données sur l'état de santé des moins de 25 ans sont l'agrégation des données granulaires d'état de santé individuel de tous les individus de moins de 25 ans
  - Les données nationales sur le nombre de lits disponibles dans les hôpitaux sont l'agrégation des données granulaires du nombre de lits disponibles dans chaque hôpital de France

# Pour chaque base nous avons passé en revue les données agrégées et granulaires à l'aune de 4 critères d'ouverture

	<u>Questions soulevées</u>	<u>Statuts possibles</u>
<b>Liberté d'accès</b>	Qui peut accéder aux données ?	<ul style="list-style-type: none"> <li>▪ <b>Accès à tous</b> : les données sont accessibles au grand public</li> <li>▪ <b>Accès restreint</b> : l'accès est autorisé seulement à certaines catégories de personnes, par exemple les chercheurs</li> <li>▪ <b>Accès fermé</b> : les données ne sont pas accessibles</li> </ul>
<b>Coût d'accès</b>	L'accès aux données est-il payant ?	<ul style="list-style-type: none"> <li>▪ <b>Gratuit</b></li> <li>▪ <b>Payant</b></li> </ul>
<b>Format de mise à disposition</b>	Les données sont-elles dans un format exploitable ?	<ul style="list-style-type: none"> <li>▪ <b>Exploitable</b> : par exemple, Excel, csv, JSON</li> <li>▪ <b>Non exploitable</b> : par exemple pdf, consultation en ligne</li> </ul>
<b>Conditions de réutilisation</b>	Les conditions générales d'utilisation des portails d'accès aux données autorisent-elles la réutilisation ?	<ul style="list-style-type: none"> <li>▪ <b>Explicité sans restriction</b> : conditions générales d'utilisation du portail d'accès aux données autorisant leur réutilisation y compris à des fins commerciales, à la condition éventuelle de mentionner la source</li> <li>▪ <b>Explicité avec restriction</b> : conditions générales d'utilisation du portail d'accès interdisant certaines réutilisations des données</li> <li>▪ <b>Non explicité</b> : absence de conditions générales d'utilisation sur le portail d'accès aux données</li> </ul>