

# UN NOUVEAU GISEMENT POUR LES STATISTICIENS ET LES ECONOMISTES : LES DONNÉES DE LA BANQUE DE FRANCE

Jacques FOURNIER, Direction Générale des Statistiques

CNIS « Les enjeux des nouvelles sources de données pour la statistique publique », 2 juillet 2018

1 - LES DONNEES DE LA BANQUE DE FRANCE

3

2 - ANONYMISATION DES DONNEES

6

3 - APPARIEMENT DES DONNEES

7

4 - APPELS A PROJETS : DEMANDE ET MODALITES D'ACCES

8

5 - L'OPEN DATA ROOM EN PRATIQUE

9

# UN OBJECTIF D'INTERET GENERAL



### Répondre au besoin de transparence du public au sens large

- sur le financement de l'économie et le fonctionnement du système bancaire et financier



### Soutenir la recherche

- française et internationale

La Banque de France a totalement ouvert ses données, dans le respect des règles juridiques (européennes essentiellement, cf. Règlement 2533-98) qui s'appliquent aux Banques Centrales :

- Stat Infos pour les informations statistiques « macro »
- Séries agrégées sur internet sur un portail statistique WEBSTAT 40 000 séries
- Données individuelles ou granulaires dans une data room OPEN DATA ROOM ; 800 millions de séries ; service public **gratuit**

*Rubrique du site internet de la Banque dédiée à l'accès aux données statistiques : [lien](#)*

## 1 - Les données de la Banque de France (2/2)



### Sources multiples

- Elles peuvent être issues de documents comptables, de statistiques internationales, d'enquêtes ou de collectes spécifiques



### Volume

- Les données accessibles portent sur les entreprises, les ménages, les établissements financiers, les titres :
- *(280 000 entreprises, environ un million de ménages et plus de 500 établissements financiers)*
- **25 milliards de données utiles**

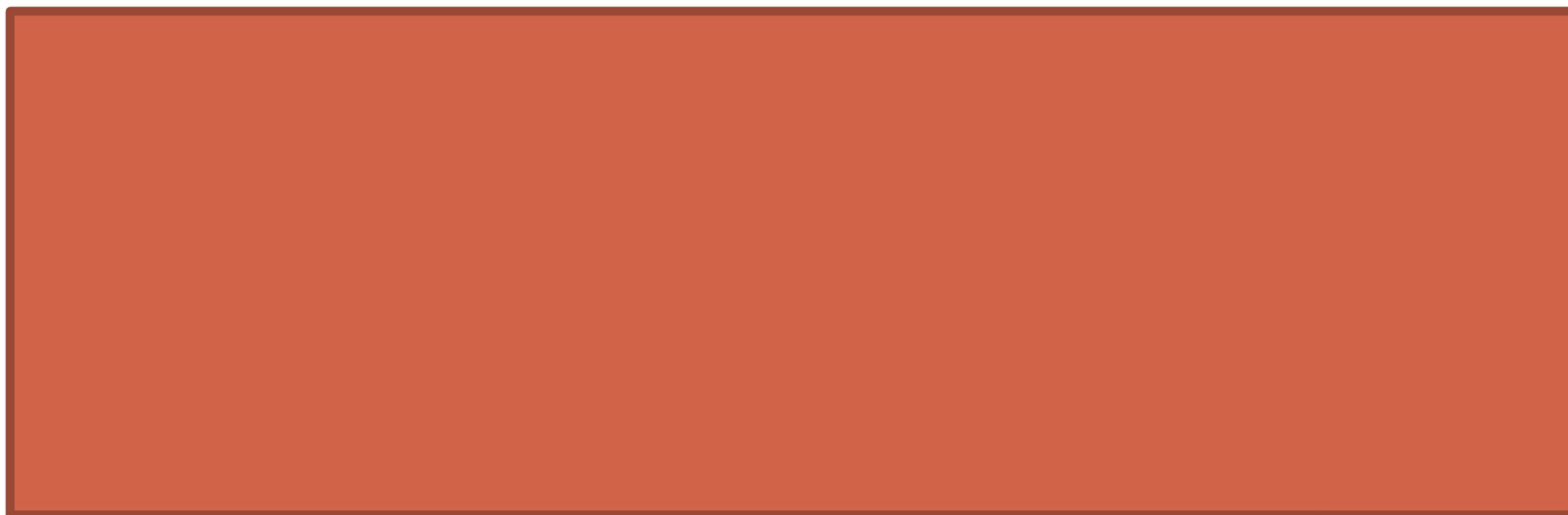


### Expertise et valeur ajoutée

- Les données sont assorties d'un dialogue avec les producteurs de données et d'un support méthodologique pendant la durée des travaux de recherche

## 2 – Anonymisation des données

- Les variables d'identification sont anonymisées pour empêcher l'identification
- La référence du projet de recherche permet de créer une clé de hachage unique par projet de recherche qui est conservée et peut être rejouée en cas d'accès ultérieur par le chercheur.



### 3 – Appariement des données

- L'appariement des données entre différents jeux de données de la Banque de France ou avec des données externes est réalisable si les bases contiennent une code d'identification commun
- Les données des différentes sources peuvent être alors anonymisées avec la même clé d'anonymisation
- Le développement de codes universels tels que le Legal Entity Identifier (LEI), soutenu par la Banque, fait écho à l'intérêt croissant des chercheurs pour les croisements de données.

LEI is the linchpin that connects the dots across the universe of entity identification in the digital age

- 20-digit, alpha-numeric code based on the ISO 17442 standard developed by the ISO



Example: ABN Sécherre SA, Switzerland



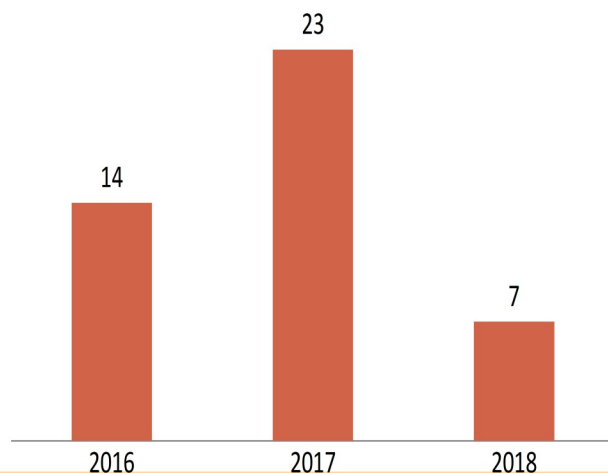
## 4 – Appels à projets : demande et procédure d'accès aux données

Demande d'accès : <https://www.banque-france.fr/statistiques/acces-aux-donnees-granulaires/open-data-room>

- Présentation synthétique du projet de recherche et des données demandées ;
- Accord du Comité d'accès aux données formé de responsables de la Banque de France et de personnalités du monde académique
- Il vérifie que les demandes correspondent à des recherches académiques, avec un projet de publication

Dépôt des demandes : [DGS-DIMOS-acces-donnees-ut@banque-france.fr](mailto:DGS-DIMOS-acces-donnees-ut@banque-france.fr)

Une cinquantaine d'équipes de recherche entre novembre 2016 et mai 2018



Quelques exemples de projets de recherche :

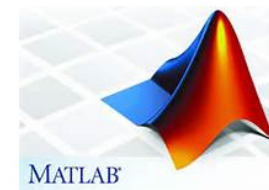
- Relation entre concurrence et innovation: impact des contraintes financières sur les entreprises françaises
- Effets de la réglementation sur le crédit bancaire alloué et sur l'investissement
- Securitization and bank liquidity management after crisis
- A general equilibrium model of asset demand and supply with heterogenous preferences...

## 5 - L'Open data room en pratique

- Les chercheurs accèderont aux données dans une salle sécurisée installée à la DGS (**Adresse : Banque de France Direction générale des statistiques 37 rue du Louvre 75002 Paris**)
- Postes de travail sécurisés, disposant d'une forte capacité de stockage et de calcul.
- Large gamme de logiciels de traitement statistique et d'optimisation :



JDEMETRA+





MERCI DE VOTRE ATTENTION

