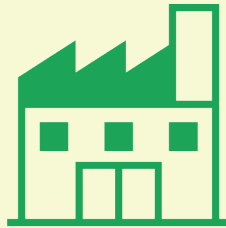




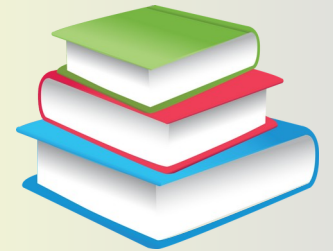
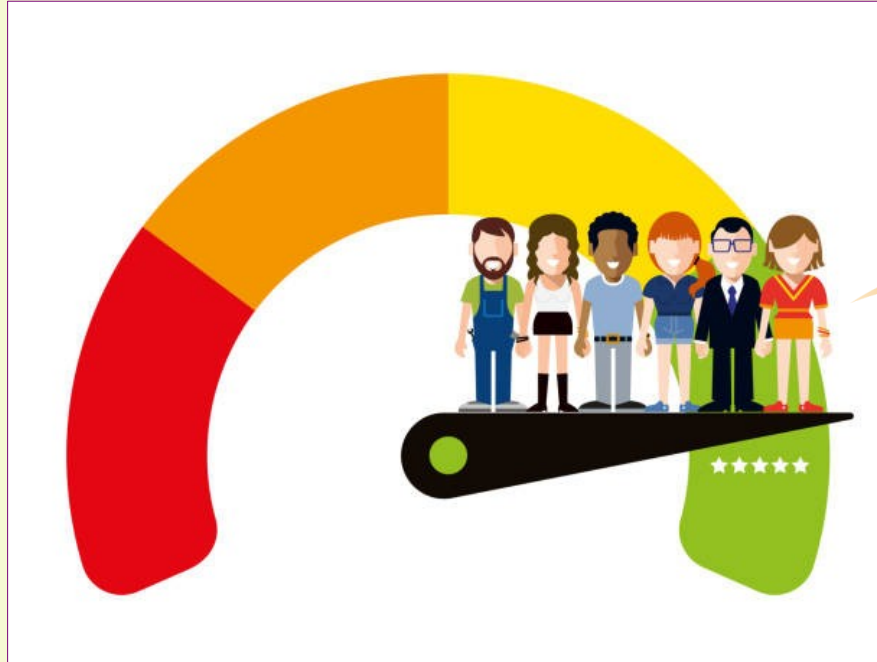
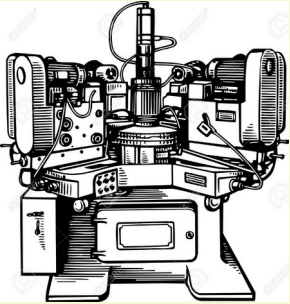
Production et diffusion des statistiques structurelles d'entreprises



Magali Demotes-Mainard
Sylvie Marchand

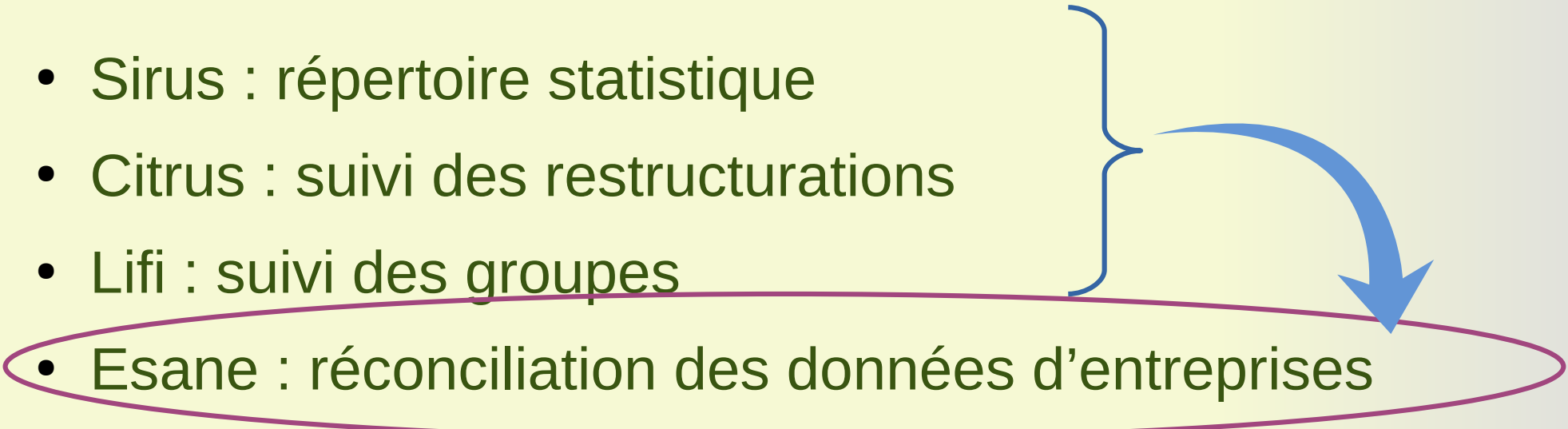
Cnis – commission Entreprises et stratégies de marché
7 octobre 2021

Les objectifs de la mission : analyser les besoins des utilisateurs



⇒ Interroger les finalités pour rétroagir sur la
« conception » du produit

Les statistiques structurelles d'entreprises

- Sirius : répertoire statistique
 - Citrus : suivi des restructurations
 - Lifi : suivi des groupes
 - Esane : réconciliation des données d'entreprises
 - Démographie d'entreprise
 - Fats : observation des filiales étrangères
- 
- A blue arrow originates from a bracket grouping the first three items (Sirius, Citrus, Lifi) and points to the 'Esane' item. The 'Esane' item is circled in purple.

Les 5 objectifs de Resane



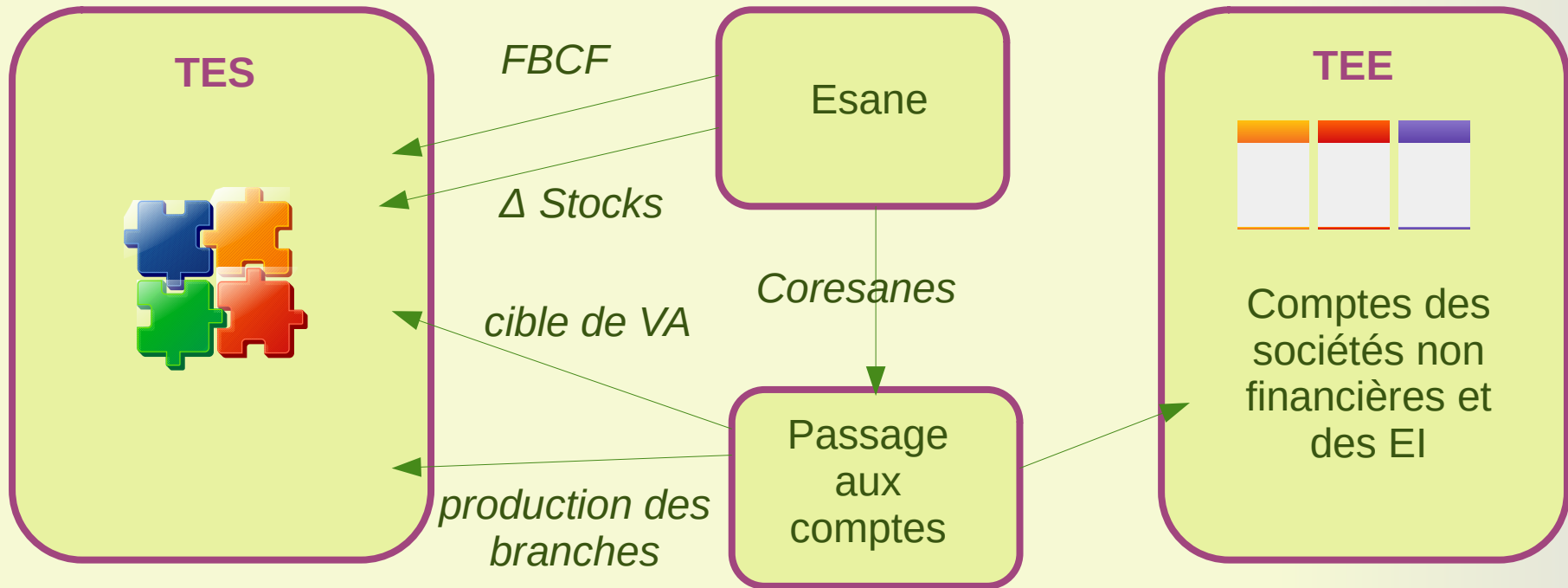
- ✓ réduire la charge statistique pesant sur les entreprises
- ✓ rationaliser les moyens consacrés à ces travaux par la statistique publique
- ✓ donner un résultat unique pour chaque variable
- ✗ raccourcir les délais de diffusion
- ✓ prendre en compte les groupes dans la statistique d'entreprise

Les idées fortes du diagnostic



- 1) La statistique d'entreprise est un domaine complexe par nature
 - *hétérogénéité de l'univers des entreprises, concepts sous-jacents aux entités étudiées et aux variables*
- 2) En matière de statistiques structurelles, l'Insee produit un ensemble de données sans équivalent
 - *en termes de nombre de variables et d'exhaustivité du champ couvert*
- 3) Ces données sont très souvent utilisées comme un « produit intermédiaire » incorporé dans d'autres traitements

Les comptes nationaux utilisent les agrégats d'Esane « en différentiel »



Les utilisateurs des données individuelles apprécient la richesse des données d'Esane



QUOI ?

QUI ?

- Des données sur longue période
- Analyser des secteurs à un niveau fin
- Des données de bonne qualité
- Apparier avec d'autres données
- Connaître la façon dont elles sont fabriquées
- Apprécier les changements de méthodes
- Une documentation structurée et rédigée
- Les entreprises mais aussi les unités légales

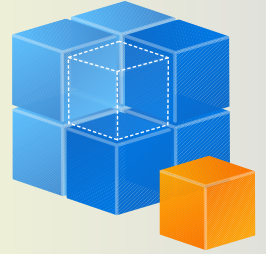
Chargés d'études, chercheurs Insee
Sectoriels Insee, SSM
Professionnels
Enseignants-chercheurs
Institutions : Trésor, DGE, Banque de France, BCE, France Stratégie...

Les idées fortes du diagnostic



- 1) La statistique d'entreprise est un domaine complexe par nature
 - *hétérogénéité de l'univers des entreprises, concepts sous-jacents aux entités étudiées et aux variables*
- 2) En matière de statistiques structurelles, l'Insee produit un ensemble de données sans équivalent
 - *en termes de nombre de variables et d'exhaustivité du champ couvert*
- 3) Ces données sont très souvent utilisées comme un « produit intermédiaire » incorporé dans d'autres traitements
- 4) Le système est arrivé à maturité

3 ensembles de recommandations



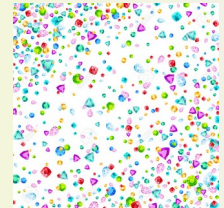
- 1) Répondre aux besoins nombreux et variés des utilisateurs
- 2) Associer les partenaires à l'évolution du dispositif Esane
- 3) Stabiliser et améliorer les méthodes de production des données d'Esane

Répondre aux besoins nombreux et variés des utilisateurs

- les « clients » de la SSE sont des utilisateurs avertis
- qui intègrent les données dans leurs propres traitements
- qui ont légitimement besoin des données en unités légales

⇒ **mettre l'usage des données individuelles en priorité des finalités d'Esane et rendre les utilisateurs autonomes**

y compris pour les unités légales



Répondre aux besoins nombreux et variés des utilisateurs

- les données d'Esane sont particulièrement riches
- les concepts et les méthodes d'élaboration de la SSE sont complexes
- certains utilisateurs sont éloignés du SSP

⇒ **documenter davantage et accompagner les livraisons de données individuelles**

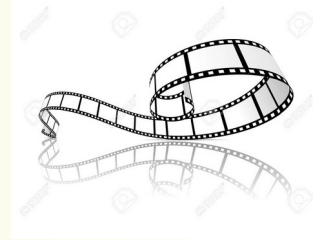
préciser et « tracer » les traitements

promouvoir le fichier Fare



Répondre aux besoins nombreux et variés des utilisateurs

- Esane, par nature, produit une photo
- de nombreux utilisateurs veulent en déduire un film
- le fichier « n-1 méthode n » ne répond pas à toutes les problématiques



⇒ **consolider et faire évoluer le dispositif permettant d'utiliser les données d'Esane en évolution**

isoler les changements de méthodes

donner des outils pour décomposer les évolutions

Associer les partenaires à l'évolution du dispositif Esane

- de nombreuses équipes impliquées dans le processus
- le ressenti d'un certain manque de collégialité



⇒ **favoriser l'expression des besoins des partenaires de premier niveau et leur donner une réponse dans des instances adaptées**

clarifier les rôles des acteurs et des instances

associer les sectoriels aux contours du profilage, selon des principes définis en commun

Associer les partenaires à l'évolution du dispositif Esane

- la Banque de France complète les SSE établies par l'Insee
 - *description des secteurs financiers, établissement des comptes financiers, élaboration de la Balance des paiements*
 - toutes ses productions restent établies sur la base de l'unité légale
 - pas de convergence organisée au niveau européen (Eurostat vs BCE)
- ⇒ engager la concertation avec la Banque de France sur les principes et les méthodes de profilage et de consolidation, pour viser une mise en œuvre cohérente

poser les jalons...



Stabiliser et améliorer les méthodes de production des données d'Esane

- un système complexe d'échantillonnage, de contrôle, d'imputation et de redressement hérité de Suse et des EAE
 - qui vise à assurer la qualité des données en entreprise profilée
 - qui privilégie l'estimation des niveaux par rapport aux évolutions
 - qui a un objectif de précision au niveau 3 positions de la NAF
- ces choix de priorité limitent certaines utilisations

⇒ **renforcer de façon ciblée les macro-contrôles d'Esane**



mieux contrôler les UL sous-jacentes aux EP, anticiper les contrôles en évolution, améliorer la précision pour certaines sous-populations pré-définies

Stabiliser et améliorer les méthodes de production des données d'Esane

- 2 méthodes de profilage
 - pour les groupes les plus complexes (cible 1) : profilage « manuel »
 - pour les autres groupes (cible 2), un algorithme qui calcule et consolide les flux intra-groupe
- les hypothèses sous-jacentes à l'algorithme ne sont pas toujours adaptées

⇒ **améliorer la méthode de profilage de la cible 2, en veillant à la cohérence avec les méthodes utilisées pour la cible 1**

tester et affiner les hypothèses

progresser dans l'observation directe des flux intra-groupe



Stabiliser et améliorer les méthodes de pondération des données d'Esane

Handwritten mathematical formulas:

$$dx^2 = dx_1^2 + dx_2^2 - dx_1 dx_2 = k^2$$
$$x = \left(\frac{u}{\sqrt{1-u^2}}, \frac{v}{\sqrt{1-v^2}} \right) \frac{m}{1-u^2} \frac{1}{\sqrt{1-u^2}}$$
$$x = \left(\frac{u_1^2 + u_2^2 + u_3^2 + u_4^2}{1-u^2}, \frac{v}{\sqrt{1-v^2}} \right) \frac{m}{1-u^2} \frac{1}{\sqrt{1-u^2}}$$
$$x = \frac{t + ux}{\sqrt{1-u^2}} \quad x = \frac{x^2 + vx}{\sqrt{1-v^2}} \quad y = y' z = z'$$

Other formulas on the right:

$$\sum_{i=1}^n \frac{1}{f_i} = \frac{2}{f_1 + f_2}$$
$$\sum_{i=1}^n \frac{1}{f_i} = \frac{2v}{f_1 + f_2}$$
$$\sum_{i=1}^n \frac{1}{f_i} = \sum_{j=1}^n \frac{1}{f_j}$$
$$\sum_{i=1}^n \frac{1}{f_i} = \sum_{j=1}^n \frac{1}{f_j}$$
$$\sum_{i=1}^n \frac{1}{f_i} = \sum_{j=1}^n \frac{1}{f_j}$$
$$\sum_{i=1}^n \frac{1}{f_i} = \sum_{j=1}^n \frac{1}{f_j}$$

- deux modes d'utilisation des données d'Esane coexistent :
 - estimateur composite ou données individuelles pondérées
 - l'estimateur composite a été introduit pour obtenir une estimation sans biais des agrégats affectés par des phénomènes décrits par les seules données d'enquête
 - mais les résultats tirés de cet estimateur sont difficiles à relier aux données individuelles, donc à interpréter (notamment en évolution)
- ⇒ mener une réflexion sur les cas d'usage de l'estimateur composite
- analyser et documenter son apport, définir ses cas d'usage*



Merci de votre attention

