Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Département fédéral de l'intérieur DFI
**Office fédéral de la statistique OFS**

Programme général de
# Statistique des entreprises

**N° 3** octobre 2010

**NEWSLETTER**
Informations sur le programme général
de statistiques des entreprises (GUS)

## Contenu

**Données secondaires**

Quels potentiels pour
les statistiques d'entreprise? 2

Avantages et inconvénients
des données secondaires 3

L'exemple des données scannées 3

## Editorial

Le programme général de statistiques des entreprises (GUS) a été initié pour moderniser le domaine de la statistique d'entreprise.

Rappelons que cette révision globale des statistiques d'entreprise vise l'utilisation accrue des données secondaires existantes, l'exploitation des potentiels de synergie entre les activités statistiques analogues, l'optimisation des processus de production statistique, et enfin une comparabilité européenne croissante sur le plan conjoncturel et structurel.

Au travers de ces objectifs, en ligne avec la politique de modernisation de la statistique fédérale selon le programme pluriannuel 2007–2011, le programme GUS s'emploie activement à satisfaire une demande croissante en information et à réduire la charge administrative qui pèse sur les entreprises enquêtées.

Ce troisième numéro présente une réflexion sur le potentiel d'utilisation des données dites «secondaires». Les données considérées comme «secondaires» pour l'OFS sont des données récoltées par des acteurs externes (administrations publiques ou entreprises privées), des données existantes qui peuvent être utilisées à des fins statistiques. Ces données se distinguent des données «primaires» qui sont des données récoltées et travaillées directement par l'OFS. Quels sont les avantages et les inconvénients de ces données? L'exemple des données scannées est présenté dans ce numéro.

La direction de programme GUS vous souhaite une lecture enrichissante.

Rick Trap, Chef du programme GUS

**Documents**

Cette Newsletter, assortie de tous les liens, est aussi proposée sur le portail Statistique suisse à l'adresse: www.statistique.admin.ch -> Actualités -> Projets de modernisation -> Programme général de statistique des entreprises (GUS) -> Publications -> Newsletter

Contacts et renseignements
Rick Trap – GUS@bfs.admin.ch

## Données secondaires:

### Quels potentiels pour les statistiques d'entreprise?

Depuis plusieurs années, les besoins en information au niveau national et international s'accroissent; l'information statistique doit être plus fréquente et plus exhaustive. Ces nouvelles exigences stimulent la recherche de nouvelles méthodes en matière de récolte de données. Paralèllement la charge administrative sera réduite et mieux répartie sur les fournisseurs de données.

Dans ce contexte, le programme pluriannuel 2007–2011 de la statistique fédérale promeut l'utilisation accrue des sources de données déjà existantes, dites «secondaires», à des fins statistiques.

L'OFS recense principalement deux types de données secondaires utiles pour la statistique d'entreprise:

- Les données administratives ou de registre qui sont des données produites par des acteurs externes dans le cadre de leur activité administrative, par exemple les données de la TVA, de l'AVS ou de douane.
- Les données produites dans le cadre d'une activité commerciale et provenant d'entreprises privées, comme les données scannées.

Les données secondaires peuvent être utilisées à des fins statistiques nécessitant la plupart du temps un processus de transformation garantissant la protection des données, avant d'être utilisée de manière optimale.

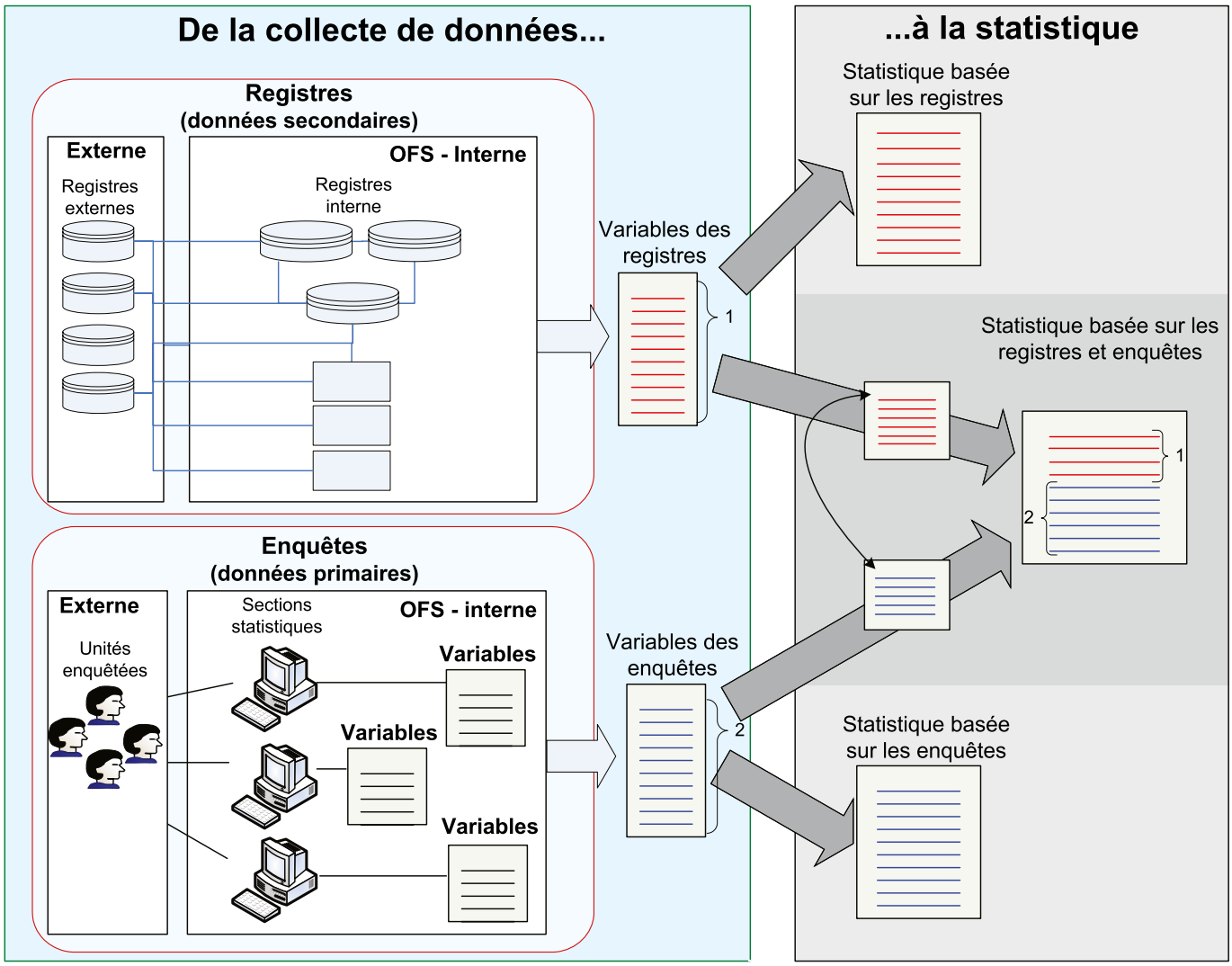
Le programme de modernisation de la statistique d'entreprise s'inscrivant dans la ligne stratégique du programme pluriannuel couvre plusieurs projets visant à étudier le potentiel de ces sources de données secondaires, notamment les deux types de données secondaires mentionnées ci-dessus.

De manière générale, on distingue deux modes d'utilisation des données administratives.

D'une part, en tant que source *directe* d'informations, permettant:

- De tirer les informations nécessaires pour une statistique directement et uniquement d'une source de données secondaires, permettant à terme de réduire le nombre de questionnaires.
- De combiner l'utilisation de données secondaires et d'enquêtes directes afin d'alléger les questionnaires statistiques en évitant de demander des informations déjà disponibles auprès d'autres sources de données.

D'autre part, en tant que source *indirecte* permettant d'améliorer l'efficacité des enquêtes et la fiabilité de leurs résultats en:




La figure ci-dessus montre l'articulation entre données d'enquêtes et données de registres.

– Tenant à jour les registres déjà existants.
– Améliorant les plans d'échantillonnage, les schémas de pondération et les méthodes d'estimation.
– Comblant les valeurs manquantes au sein d'enquête de petite taille.

L'utilisation des sources de données secondaires est déjà une réalité pour certaines enquêtes, notamment avec les données de la TVA ou les données scannées.

## Avantages et inconvénients de l'utilisation de données secondaires

**Avantages**

De manière générale, l'utilisation des données secondaires à des fins statistiques représente les avantages suivants:

– Les données secondaires sont recueillies régulièrement avec une relativement bonne qualité, ce qui a un effet positif sur leur utilisation dans le processus de production statistique.
– La récolte continue des données secondaires permet de produire, dans la plupart des cas, des informations statistiques à un rythme accéléré.
– Les données du registre représentent un grand potentiel pour le développement de nouveaux types de statistiques. Grâce au processus même d'actualisation du registre, les données longitudinales sont très riches et permettent de suivre les unités statistiques à travers le temps.
– Les données secondaires servent à compléter, corriger et à améliorer la qualité des données d'enquêtes.
– L'utilisation optimale des données secondaires permettra à terme de diminuer les coûts d'enquêtes directes et à soulager nettement la charge administrative des entreprises.
– Les données secondaires permettent une couverture totale des unités observées d'un univers donné.

**Inconvénients**

Les sources secondaires, ayant des objectifs divergents de ceux de la statistique, représentent également quelques limites pour le système statistique:

– Les définitions des unités et des variables récoltées par des acteurs externes ne correspondent pas toujours aux besoins statistiques et doivent être remaniées afin d'être compatibles avec le système statistique.
– Certaines variables spécifiques obtenues auparavant par les recensements ou sondages ne sont plus disponibles. Des solutions d'enquêtes par sondage sur des sous-groupes de population réduits restent nécessaires pour couvrir les besoins des utilisateurs et producteurs de statistiques.
– Tout changement touchant les registres administratifs ou données d'entreprises peut avoir une forte influence sur les processus statistiques.
– L'investissement initial pour l'implémentation des ces nouvelles méthodes demeure important.

## L'exemple des données scannées

### Depuis la caisse enregistreuse jusque dans l'indice des prix à la consommation: la modernisation du relevé des prix

L'exploitation des données scannées est un exemple d'utilisation des données secondaires par l'Office fédéral de la statistique (OFS).

Les codes-barres de millions d'articles sont scannés quotidiennement aux caisses des grands distributeurs. Les «données scannées» qui en résultent contiennent une grande quantité d'informations offrant de nouvelles possibilités pour le relevé des prix dans le cadre de l'établissement de l'indice suisse des prix à la consommation (IPC) et de l'indice des prix à la consommation harmonisé (IPCH).

L'IPC mesure le renchérissement des marchandises et des services représentatifs de la consommation des ménages privés. Établi chaque mois par l'OFS, il est utilisé par un vaste public. Il sert notamment à l'indexation des salaires et des rentes, à la préparation des bases de décision de politique économique (par exemple, pour la Banque nationale) et à l'apurement des données statistiques sur les prix (par exemple, les Comptes nationaux). L'IPCH, également établi par l'OFS depuis 2008, se base principalement sur les mêmes données concernant les prix que l'IPC et sert en premier lieu à comparer le renchérissement observé en Suisse avec celui dans d'autres pays d'Europe.

L'OFS et un institut d'étude de marché mandaté à cet effet relèvent chaque mois quelque 50'000 prix en Suisse pour calculer l'IPC et l'IPCH. Pour la plupart des articles, le relevé est réalisé directement dans les points de vente, par téléphone, par Internet ou par courrier postal.

Depuis juillet 2008, l'OFS utilise non plus seulement des prix relevés de manière traditionnelle mais aussi des données scannées pour établir l'IPC et l'IPCH. Ces dernières comportent des informations détaillées sur les prix et les chiffres d'affaires et sont en outre disponibles peu de temps après avoir été enregistrées à la caisse. Leur utilisation pour le calcul de l'indice présente de nombreux avantages par rapport au relevé traditionnel des prix dans les points de vente.

Compte tenu de la quantité d'informations qu'elles contiennent et de leur actualité, les données scannées permettent d'obtenir des résultats d'une plus grande qualité grâce à une plus large couverture temporelle et spatiale des données sur les prix. Alors que les prix étaient relevés traditionnellement une fois au cours des six premiers jours ouvrables de chaque mois, les données scannées peuvent l'être sans interruption pendant les 14 premiers jours du mois. Les variations de prix survenues avant et après l'ancien jour de relevé sont ainsi également prises en compte, de même que, pour chaque article, le prix payé par chacun des consommateurs de tous les points de vente du grand distributeur. Les informations sur les chiffres d'affaires relevées dans les données scannées permettent par ailleurs de sélectionner précisément le produit le plus représentatif pour

le relevé des prix. Le recours aux données scannées a par ailleurs pour avantage de décharger les grands distributeurs de certaines tâches administratives. Lors du relevé traditionnel des prix, les enquêteurs et enquêtrices devaient régulièrement faire appel au personnel de vente afin de pouvoir définir les produits générant le plus de chiffre d'affaires et les changements dans l'assortiment. Tout ce travail n'est plus nécessaire avec la transmission systématique des données scannées.

Les données scannées permettent par ailleurs à moyen terme de réduire les coûts du relevé des prix. Le potentiel est le plus important pour les nombreux produits frais (alimentation), dont les prix doivent être relevés séparément dans diverses régions.

Les expériences faites jusqu'ici à l'OFS et à l'étranger ont montré que la collecte de ces données est liée à des défis importants. Ces dernières ne sont introduites que progressivement pour le relevé des prix dans le cadre de l'IPC et de l'IPCH. Certains groupes de marchandises particulièrement difficiles à traiter (non food) sont encore exclus. La première étape a débuté en juillet 2008 avec l'introduction des données scannées d'un grand distributeur pour les groupes de marchandises «alimentation, produits à lessive et de nettoyage» et «produits cosmétiques». Depuis avril 2010, ces mêmes données d'un autre grand distributeur sont aussi prises en considération.

Il est prévu d'étendre progressivement le recours aux données scannées à d'autres grands distributeurs et à d'autres groupes de marchandises au fil des prochaines années. Il s'agira aussi d'évaluer d'autres méthodes de relevé et de calcul permettant d'exploiter encore davantage le potentiel d'amélioration et de rationalisation des données scannées.

**Impressum**

La newsletter peut être obtenue sous forme électronique dans le cadre d'un abonnement gratuit. Si vous souhaitez recevoir la newsletter, nous vous prions de vous inscrire à l'adresse suivante: www.news-stat.admin.ch.
Son but est d'informer les clients de la statistique des entreprises, les personnes intéressées, les entreprises, les services publics et d'autres institutions. Les articles signés exposent l'avis de leurs auteurs.

**Editeur**
Office fédéral de la statistique
Espace de l'Europe 10
2010 Neuchâtel

**Renseignements et commandes**
E-mail: gus@bfs.admin.ch

**Layout et graphisme**
Section DIAM/PP

**Traduction**
Service de traduction de l'OFS

**Numéro de commande**
903-1001