



Statistique démographique longitudinale (DVS)

Règles de production et résultats

Neuchâtel, 2022

Éditeur:	Office fédéral de la statistique (OFS)	Concept de mise en page:	Section PUB
Renseignements:	info.dem@bfs.admin.ch	Images:	Section POP
Contenu:	Section POP	Téléchargement:	www.statistique.ch
Domaine:	01 Population	Copyright:	OFS, Neuchâtel 2022
Langue du texte original:	Français		La reproduction est autorisée, sauf à des fins commerciales, si la source est mentionnée

Versions

<i>Version</i>	<i>Date</i>	<i>Description</i>	<i>Nom</i>
01	20.08.2020	Version initiale du module Migrations	Cri, JUSA, KAH
02	30.12.2020	Mise à jour avec le module Nationalité	Cri, KAH, JUSA
03	30.06.2021	Mise à jour suite aux modifications des programmes	KAH, Cri, JUSA, Sa
04	09.09.2021	Intégration des remarques de B. Weber (SECO)	Cri
05	28.09.2021	Mise à jour des chiffres	KAH
06	15.06.2022	Mise à jour avec le module Autorisation de résidence	Cri, JUSA, KAH

Liste des abréviations

<i>Abréviation</i>	<i>Signification</i>
BEVNAT	Statistique des mouvements naturels de la population
CdC	Centrale de compensation (organe central d'exécution de la Confédération en matière d'assurances sociales du 1 ^{er} pilier)
KD	Banque de données consolidées de l'OFS (Konsolidierte Datenbank)
Rafixe	Fichier mis à disposition par la CdC permettant de mettre à jour les NAVS13 à une date de référence et d'identifier les NAVS13 annulés
NAVS13	Numéro d'assuré à 13 positions
OFS	Office fédéral de la statistique
ORDIPRO	Système d'information destiné à la saisie des données relatives aux diplomates étrangers et fonctionnaires internationaux, exploité par le Département fédéral des affaires étrangères (DFAE)
STATPOP	Statistique de la population et des ménages
SYMICAfile	Système d'information central sur la migration (pour les personnes en processus d'asile)
SYMICLEtr	Système d'information central sur la migration (pour les personnes soumises à la loi sur les étrangers)
UPI	« Unique Person Identification » : fonctionnalité du Registre Central des Assurés des assurances sociales fédérales qui implémente l'identification administrative de personnes physiques et la gestion de l'identificateur NAVS13

Table des matières

1	But du document	4
2	Situation actuelle	4
3	Sources de données utilisées	4
4	Développement modulaire	5
5	Préparation des données longitudinales	5
5.1	Univers de la population.....	6
5.2	Harmonisation des fichiers	6
5.3	Numéros AVS	7
5.4	Table des attributs	7
5.5	Création de mouvements	8
6	Biographies cohérentes	9
6.1	Création des biographies	9
6.2	Traitement des biographies incohérentes.....	10
6.2.1	Traitement des immigrations et émigrations	10
6.2.2	Traitement des nationalités	11
6.2.3	Traitements des autorisations de résidence	12
6.2.4	Traitement du type de population	13
6.3	Flags	13
7	Fichiers mis à disposition	14
8	Résultats	14
8.1	Résultats généraux	14
8.2	Valeurs manquantes	15
8.3	Comparaison des résultats avec STATPOP	15
8.4	Données rétroactives.....	19
	Annexes	21

1 But du document

Ce document décrit les règles de production et montre les résultats de la statistique démographique longitudinale. Il doit permettre de comprendre comment est créé le produit longitudinal, dynamique et cohérent dans le temps, basé sur les données de la statistique de la population et des ménages (STATPOP).

2 Situation actuelle

Avec l'introduction du recensement annuel en 2010 ainsi que l'utilisation systématique du numéro d'assuré (NAVS13) dans les données, le potentiel des analyses longitudinales dans le domaine de la population est devenu énorme.

Jusqu'à présent, la production de STATPOP se limitait à la production des effectifs ou des mouvements trimestriels et annuels, sans tenir compte des données des années précédentes. Dans ce produit longitudinal basé sur les données de production et statistiques de STATPOP, tous les résultats sont recalculés chaque année. Les événements annoncés tardivement sont ainsi intégrés rétroactivement dans les années précédentes sans limite de temps et les variables qui ne changent pas dans le temps, comme par exemple la date de naissance ou le lieu de naissance, sont systématiquement contrôlées. Les résultats de cette statistique longitudinale ne sont, par conséquent, pas exactement identiques aux données publiées pour STATPOP (voir le chapitre 8.3 pour les détails sur les différences de résultats).

Ce nouveau produit permet à l'OFS de calculer et de mettre à jour facilement des indicateurs longitudinaux dans le domaine de la population, mais il permet surtout de proposer aux utilisateurs internes et externes des données exhaustives de la population, harmonisées à une date de référence et plausibilisées dans le temps au niveau individuel. Chaque personne présente dans la statistique a une biographie cohérente. Le suivi de cohortes est ainsi directement possible, sans travaux de préparation de données préalables.

3 Sources de données utilisées

La statistique démographique longitudinale se base principalement sur des micro-données consolidées de STATPOP (effectifs trimestriels, effectifs annuels, mouvements). Ces données sont directement lues dans la banque de données consolidée (KD) de l'OFS, où elles sont stockées sans noms, prénoms ou adresses et dont les identificateurs sont pseudonymisés.

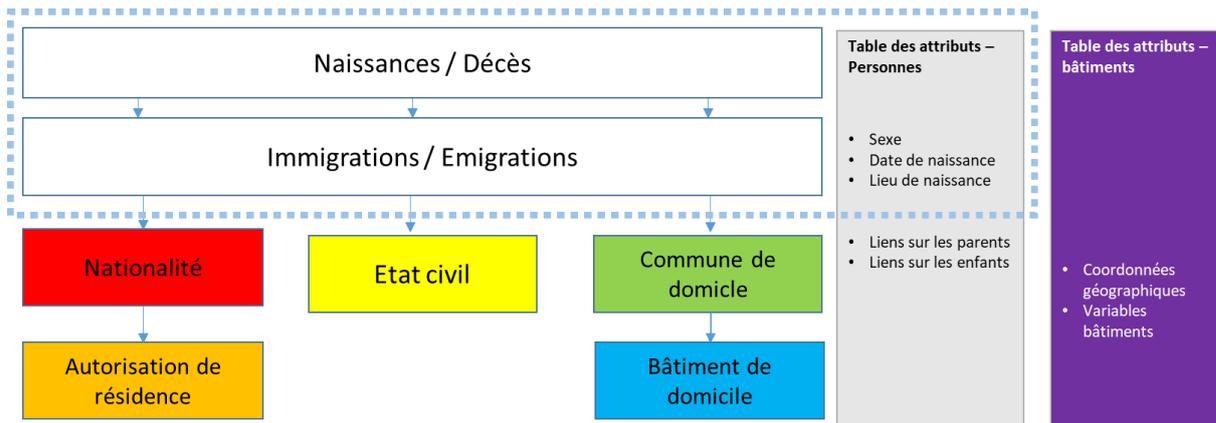
D'autres données livrées ou produites dans le cadre de la production STATPOP sont aussi utilisées pour la production de la statistique longitudinale. Il s'agit d'un extrait du registre UPI (Unique Person Identification) géré par la caisse de compensation (CdC), du fichier « Rafixe » de la CdC qui permet de mettre à jour les NAVS13 qui auraient changé ainsi que de la table des enregistrements qui est la mise en commun trimestrielle de tous les enregistrements reçus des registres des habitants (RdH), du système d'information central sur la migration (SYMIC) et du registre des diplomates et fonctionnaires internationaux (ORDIPRO).

Contrairement à STATPOP dans laquelle les bilans utilisent directement les données de la statistique des mouvements naturels de la population (BEVNAT), dans la statistique longitudinale, les naissances, les décès ainsi que les changements d'états civils sont recréés à partir des informations reçues des RdH, de SYMIC et d'ORDIPRO. Le but étant d'éviter de créer des biographies pour des personnes qui n'auraient aucune apparition dans STATPOP, mais qui auraient un événement dans BEVNAT.

4 Développement modulaire

La démarche que nous suivons pour le développement de la statistique longitudinale est de type modulaire : un produit de base, appelé module « Migrations », contenant les immigrations, les émigrations, les naissances et les décès ainsi que les variables correspondantes est développé dans un premier temps (produit blanc). Chaque thématique supplémentaire est développée de façon semi-indépendante (en lien seulement avec les modules situés au-dessus, selon le schéma 1) et intégrée au système au fur et à mesure de sa disponibilité.

Schéma 1 : Développement par module thématique

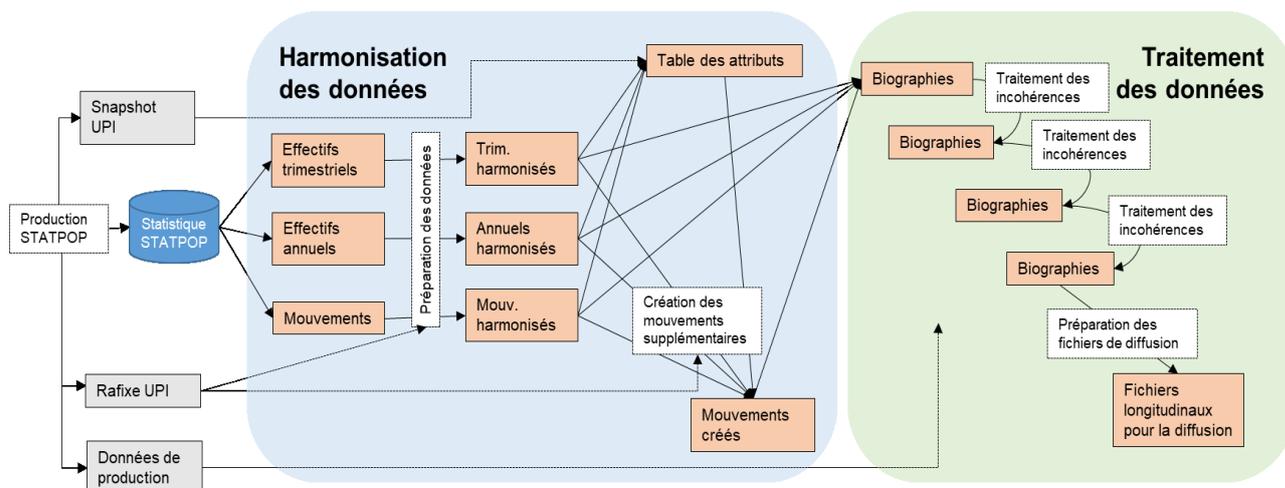


Les modules « Migrations » (en pointillé bleu), « Nationalité » et « Autorisation de résidence » sont disponibles. La manière dont ces données sont produites est décrite dans la suite de ce document, dans laquelle l'ensemble des trois modules est appelé « module Autorisation de résidence ». Avec l'ajout du module « Autorisation de résidence », la variable « Type de population » qui permet de distinguer les enregistrements de la population résidante permanente de la population résidante non permanente a pu être calculée. Les informations sur les liens entre parents et enfants ont aussi été ajoutés. Dans l'annexe 1 se trouve la liste des variables de base de STATPOP qui seront intégrées dans la statistique longitudinale et le module dans lequel elles sont prévues.

5 Préparation des données longitudinales

Comme illustré par le schéma 2, la production des données longitudinales se déroule en deux grandes étapes. La première (en bleu) consiste à préparer et harmoniser les données et la seconde (en vert) à plausibiliser les biographies et à traiter leurs potentielles incohérences.

Schéma 2 : production des données longitudinales



5.1 Univers de la population

Toutes les personnes qui apparaissent au moins une fois dans la population résidante permanente ou non permanente d'un effectif de STATPOP (trimestriel ou annuel) ou qui ont un mouvement d'arrivée en Suisse (immigration) ou de départ de Suisse (émigration) dans STATPOP font partie de l'univers de base¹.

Le début de la statistique longitudinale est fixé au 31.12.2010 et elle se termine avec les dernières données STATPOP disponibles.

5.2 Harmonisation des fichiers

Une sélection de variables STATPOP est faite dans la KD pour les types d'enregistrements suivants :

- Effectifs annuels de 2010 à AAAA (où AAAA = la dernière année disponible)
- Effectifs trimestriels du 31.12.2010 au 31.12.AAAA
- Mouvements annuels de 2011 à AAAA :
 - Immigrations
 - Emigrations
 - Acquisitions de la nationalité suisse
 - Changements de commune

Les variables "communes" et "pays", comme par exemple la commune de naissance ou le pays de provenance, sont systématiquement actualisées sur la base de la nomenclature valable au jour de référence du dernier STATPOP pris en compte (= jour de référence de la statistique longitudinale). Par conséquent, une personne qui habite dans une commune qui a fusionné il y a quelques années, se trouvera toujours dans la même commune dans la statistique longitudinale même si la date à laquelle se réfère l'analyse se situe avant la fusion.

¹ Il est possible qu'un nourrisson décédé dans ses premières semaines de vie ne figure pas dans l'univers de la statistique longitudinale.

Les imputations statistiques qui ont été faites dans STATPOP sur le lieu de naissance (Suisse / étranger) et sur le lieu de provenance (Suisse / étranger) ne sont pas reprises dans les données harmonisées.

5.3 Numéros AVS

La première introduction du NAVS13 dans les registres qui livrent leurs données pour la production de STATPOP a été réalisée en 2009, c'est à ce moment-là que le registre UPI, qui permet le contrôle systématique de l'attribution des NAVS13 de nos enregistrements, a été créé. Etant donnée la jeunesse du système en 2010, il a été décidé en 2014, de rechercher à nouveau les NAVS13 des données du 31.12.2010 sur la base des noms et prénoms. Le résultat de cette recherche a été utilisé pour augmenter le taux d'attribution des NAVS13 des données STATPOP 2010 pour la statistique longitudinale.

Pour gagner encore quelques NAVS13, les numéros d'identification du registre SYMIC (ZEMIS_Nr) ont été utilisés. Si deux enregistrements possèdent le même SYMIC_Nr mais que seul un enregistrement a un NAVS13 valide, le NAVS13 valide est repris pour le deuxième enregistrement.

Une fois ces traitements spécifiques effectués, les numéros AVS de tous les enregistrements sélectionnés dans la KD sont mis à jour à l'état du jour de référence de la statistique longitudinale. Sur la base du registre UPI, la CdC met à disposition un fichier (« rafixe ») contenant l'ensemble des numéros d'assuré ayant existé. Pour chaque numéro listé, figure le NAVS13 correspondant valide pour la date de référence du fichier. Pour la très grande majorité des cas, le numéro de « base » et le numéro « final » sont identiques, mais il arrive que deux numéros attribués initialement concernaient, après vérification, une même personne et ont été fusionnés. Dès lors, les deux numéros initiaux correspondent au même numéro final (un des deux initiaux).

La dernière étape de traitement des NAVS13 consiste à supprimer les doublons qui ont été créés avec la mise à jour des numéros. Seuls les enregistrements avec un NAVS13 valide sont conservés pour la suite de la production des données longitudinales.

5.4 Table des attributs

Pour l'ensemble des personnes de l'univers de la population considéré pour la statistique longitudinale, une entrée dans la table des attributs est créée avec le NAVS13. Pour chaque entrée, les variables : sexe, date de naissance, pays et commune de naissance sont complétées avec l'information connue de l'enregistrement le plus récent trouvé pour la personne dans le pot de données harmonisées.

Les dates de décès sont ensuite ajoutées pour toutes les personnes se trouvant dans la table des attributs et pour lesquelles une date de décès existe dans les données de l'UPI valables au jour de référence. D'autres variables seront ajoutées dans la table des attributs au fur et à mesure du développement de la statistique longitudinale. Actuellement les variables disponibles sont notamment la date d'acquisition de la nationalité suisse, la nationalité et l'autorisation de résidence les plus récentes.

Seuls les mouvements harmonisés, qui ont une date de mouvement (dateMove) entre la date de naissance et la date de décès définies dans la table des attributs, sont conservés pour la suite de la production de la statistique longitudinale.

Les enregistrements de la table des attributs sont aussi complétés par les numéros d'assurés (NAVS13) et le sexe des parents et les numéros d'assurés des enfants pour lesquels l'information du lien de filiation a pu être reprise des données d'Infostar livrées pour la production STATPOP, si elle répond aux règles suivantes :

- Il y a plus de 12 ans de différence entre la date de naissance d'un parent et celle d'un enfant (si une des deux personnes concernées ne se trouve pas dans STATPOP, la date de naissance de l'UPI est utilisée pour calculer la différence d'âge) ;
- Chaque enfant a au maximum deux parents (ils peuvent être de même sexe) ;
- En cas de contradiction entre les enregistrements du parent et de l'enfant, ce sont les informations dans l'enregistrement de l'enfant qui sont favorisées ;
- En cas de contradiction entre les différents enregistrements d'une personne, ce sont les informations dans l'enregistrement le plus récent qui sont favorisées.

5.5 Création de mouvements

A côté des mouvements STATPOP repris de la KD : immigrations, émigrations, acquisitions de la nationalité suisse et changements de commune, d'autres mouvements sont créés sur la base de l'ensemble des données harmonisées. Il s'agit :

- *Des naissances* : pour chaque personne sélectionnée dans l'univers de la statistique longitudinale qui a une date de naissance après le 31.12.2010 et qui n'a, ni présence en 2010, ni une immigration comme premier mouvement, un mouvement de naissance est créé.
- *Des décès* : pour chaque personne sélectionnée dans l'univers de la statistique longitudinale qui a une date de décès après le 31.12.2010 dans la table des attributs et dont le dernier mouvement n'est pas une émigration, un mouvement de décès est créé avec la date de décès comme date de mouvement.
- *Des annulations de numéros AVS* : pour chaque personne sélectionnée dans l'univers de la statistique longitudinale qui a une date d'annulation du NAVS13 dans le fichier « rafixe », un mouvement d'annulation du NAVS13 est créé avec la date d'annulation comme date de mouvement. Dans la gestion des NAVS13, il est possible que des numéros AVS soient annulés. C'est par exemple le cas s'il s'avère qu'un numéro d'assuré ait été attribué à deux personnes distinctes. Dans ce cas, le NAVS13 initial est annulé et deux numéros différents sont attribués. Ces annulations de numéros d'assuré ont aussi lieu dans des situations de témoins ou de victimes à protéger ou plus fréquemment dans les cas d'adoptions. Au moment où la personne change d'identité, le numéro initial est annulé et un nouveau numéro est attribué. Dans tous les cas d'annulation, le lien entre la personne avant et après l'annulation n'est pas possible.
- *Des changements d'autorisation de résidence* : pour chaque personne sélectionnée dans l'univers de la statistique longitudinale et pour laquelle il y a une différence d'autorisation de résidence entre deux effectifs trimestriels STATPOP successifs, un mouvement de changement d'autorisation de résidence est créé le premier jour du trimestre lors duquel le changement a eu lieu. Les changements d'autorisation de résidence créés dans la statistique longitudinale concernent une agrégation des statuts disponibles dans STATPOP. Seules les valeurs suivantes sont disponibles :
 - Livret B = 2 → codes STATPOP : 02, 0201, 0202 ;
 - Livret C = 3 → codes STATPOP : 03, 0301, 0302 ;
 - Livret Ci = 4 → codes STATPOP : 0401, 0402 ;
 - Livret F = 5 → code STATPOP : 0503 ;
 - Livret L avec une autorisation de plus de 12 mois = 712 → codes STATPOP : 070101, 070201, 070907 ;
 - Livret L avec une autorisation n'atteignant pas 12 mois = 704 → codes STATPOP : 07, 0701, 0702, 070102, 070202, 070103, 070104, 070204, 070105, 070205, 070206 ;
 - Livret N = 8 → code STATPOP : 0804 ;
 - Autorisation de résidence pour diplomates et fonctionnaires internationaux avec immunité diplomatique = 11 ;

- Autorisation de résidence pour fonctionnaires internationaux sans immunité diplomatique = 12 ;
- Autorisation de résidence pas attribuée = 13.
- *Des changements de type de population* : sur la base des changements d'autorisation de résidence et des dates d'arrivée en Suisse, les mouvements de changements de type de population sont créés. Pour tous les changements d'autorisation de résidence qui permettent de changer de type de population (p.ex. passage d'un livret L < 12 mois à un livret L > 12 mois), un changement de type de population est associé au changement d'autorisation de résidence. Pour les personnes du domaine de l'asile, un changement de type de population est créé un an après leur arrivée en Suisse.

D'autres mouvements de base seront créés au fur et à mesure du développement de la statistique longitudinale.

6 Biographies cohérentes

6.1 Création des biographies

Une fois les données de base disponibles et harmonisées, l'effectif 2010 est ajusté avec les informations arrivées dans des livraisons plus récentes et qui ont un impact sur 2010. Il s'agit d'une part d'ajouter dans l'effectif 2010, les personnes qui ont une date arrivée avant le 01.01.2011 mais qui ne se trouvent pas dans l'effectif 2010 et d'autre part de supprimer les personnes présentes dans l'effectif 2010 mais qui ont une annonce de départ avant le 01.01.2011.

Tous les enregistrements pertinents de chaque personne sont ensuite regroupés et mis dans l'ordre chronologique du plus ancien au plus récent. Pour le module « Migrations », les types d'enregistrements suivants sont considérés :

- Les présences dans l'effectif 2010
- Les naissances
- Les immigrations
- Les décès
- Les émigrations
- Les annulations de NAVS13
- Les présences dans l'effectif à la date de référence des données longitudinales

Les types d'enregistrements a à c sont considérés comme des entrées dans la statistique, les types d'enregistrements de d à g sont considérés comme des sorties de la statistique. Toutes les incohérences d'une biographie sont repérées et traitées selon un système de règles déterministes définies (résumées dans le chapitre 6.2), de manière à ce que les biographies finales de toutes les personnes soient plausibles et cohérentes.

En tenant compte du fait que seuls les événements qui ont lieu lorsque la personne est en Suisse sont pertinents, les règles de plausibilité des biographies définies pour les modules « Migrations », « Nationalité » et « Autorisation de résidence » sont les suivantes :

Migrations :

- Toutes les biographies doivent commencer par une « entrée » ;
- Toutes les biographies doivent se terminer par une « sortie » ;
- Aucun événement ne peut se produire avant une naissance ;
- Aucun événement ne peut se produire après un décès ou une annulation de NAVS13 ;
- Il n'est pas possible d'avoir deux immigrations ou deux émigrations de suite ;

- Après une émigration, le seul événement possible est une immigration ;
- Une immigration ne peut pas être suivie d'une émigration moins de 4 jours plus tard ;
- Deux émigrations ne peuvent pas avoir lieu en moins de 4 jours ;
- Une personne ne peut pas avoir plus de 12 immigrations sur une même année.

Nationalité :

- Au maximum une nationalité étrangère est considérée pour une personne ;
- Le seul cas pour lequel une personne peut avoir deux nationalités différentes dans sa biographie est celui d'une acquisition de la nationalité suisse ;
- Le passage d'une nationalité étrangère à la nationalité suisse ne peut être expliqué que par la présence d'une acquisition ;
- Une acquisition est comptée comme un mouvement, seulement si la personne était en Suisse lors de l'acquisition.

Autorisation de résidence:

- Les personnes étrangères doivent avoir une autorisation de résidence ;
- Les Suisses ne doivent pas avoir d'autorisation de résidence ;
- Les nouveau-nés en attente d'une constatation de la nationalité suisse ont une autorisation de résidence inconnue (-9) ;
- Une personne ne peut pas avoir d'autorisation de résidence non attribuée (code 13) entre deux autorisation de résidence valables.
- Les enfants de moins de 15 ans ne peuvent pas avoir de livret Ci (pour conjoint ou enfant avec activité lucrative de membres de missions diplomatiques ou de fonctionnaires d'organisations internationales).

Type de population :

- Les Suisses, les diplomates et fonctionnaires internationaux, les personnes avec un livret B, un livret C, un livret Ci ou un livret L permettant de rester 12 mois en Suisse font partie de la population résidante permanente ;
- Les étrangers sans indication d'autorisation de résidence (-9), avec une autorisation de résidence pas attribuée (13) ou avec un livret L ne permettant pas de rester 12 mois en Suisse font partie de la population résidante non permanente ;
- Les personnes en procédure d'asile (livret F et N) font partie de la population résidante non permanente à leur arrivée et intègre la population résidante permanente après 12 mois de présence en Suisse.
- Les nouveau-nés dans le domaine de l'asile qui naissent en Suisse reprennent le type de population de leur mère si le lien est connu. Si le lien n'est pas connu, ils restent dans la population résidante non permanente durant les 12 premiers mois de leur vie.

6.2 Traitement des biographies incohérentes

Lorsque des incohérences sont identifiées dans les biographies, elles sont traitées selon des règles déterministes prédéfinies.

6.2.1 Traitement des immigrations et émigrations

Lorsque l'arrivée d'une personne dans l'univers considéré n'est expliquée ni par une naissance ou une immigration, ni par une présence en 2010, un mouvement d'immigration est alors imputé. Le choix de la date de l'immigration se fait sur plusieurs critères : en premier lieu, la date de la dernière arrivée en Suisse « CHarrivalDate » présente dans les effectifs ou les mouvements est reprise, si elle n'est pas disponible, la date d'arrivée dans la commune « arrivalDate » est reprise. Si aucune des deux dates n'est renseignée dans aucun effectif ou mouvement, la date d'immigration est imputée au 1^{er} janvier de

la première année où la personne est annoncée dans une commune. Si ce n'est pas possible, la date d'immigration est imputée le lendemain de la dernière émigration.

Lorsque le départ d'une personne dans l'univers considéré n'est justifié ni par un décès, une émigration ou une annulation de NAVS13, ni par une présence dans le dernier effectif de STATPOP pris en compte, une émigration est imputée. La date de mouvement de l'émigration imputée est fixée au 31 décembre de l'année précédant la dernière année où la personne est annoncée dans une commune. Par exemple, si une personne apparaît dans les données STATPOP du 31.12.2016, la date d'émigration sera fixée au 31.12.2015. Dans l'ordonnance sur l'harmonisation des registres, les contrôles des habitants des communes et des cantons doivent livrer à l'OFS toutes les personnes avec un domicile dans la commune le jour de référence ainsi que les personnes qui sont parties de la commune ou décédées dans les 12 mois précédant le jour de référence. Il arrive régulièrement que des personnes parties à l'étranger n'annoncent pas leur départ dans leur dernière commune de domicile et que l'administration réalise plus de 12 mois plus tard que la personne est partie. Ce départ saisi dans le registre des habitants n'est jamais transmis à l'OFS et n'existe donc pas dans les émigrations de STATPOP.

Les biographies avec plus de 12 immigrations sur une même année se voient supprimer les immigrations et émigrations de l'année en question. Les mouvements de ces personnes ayant lieu sur les années précédentes et suivantes sont conservés pour ne pas « perdre » leur biographie. L'année ciblée semble être le reflet de problèmes administratifs car il est improbable qu'une personne soit enregistrée dans un registre avec un domicile en Suisse, plus de 12 fois sur une année.

Les immigrations suivies d'une émigration dans les trois jours suivants sont aussi traitées car elles sont certainement dues à des problèmes administratifs. Il est improbable qu'une personne soit réellement venue en Suisse et inscrite dans un registre des habitants que pour quelques jours. Par contre, les émigrations suivies d'une immigration dans les trois jours qui suivent sont conservées. Il s'agit en grande majorité de renouvellements de livret L lors desquels la personne est censée repartir à l'étranger. Par contre, les biographies contenant deux émigrations en l'espace de 3 jours sont traitées, car il s'agit vraisemblablement aussi d'un problème administratif, soit lors d'un renouvellement de livret ou lors d'un départ de la Suisse, dans ce cas une seule émigration est conservée.

Les émigrations sans lieu de destination suivies d'une immigration ou un déménagement dans les 14 jours ont été supprimées (« fausses » émigrations). Il ne s'agit probablement pas d'un départ à l'étranger mais plutôt d'un déménagement en Suisse. De plus, les immigrations sans lieux de provenance faisant suite à une « fausse » émigration ont aussi été supprimées (« fausses » immigrations).

6.2.2 Traitement des nationalités

De manière générale, la nationalité annoncée dans les données les plus récentes sera toujours prise en compte.

Si une personne a une nationalité étrangère dans sa dernière présence dans les données STATPOP (effectif trimestriel, effectif annuel ou mouvement), alors cette nationalité sera considérée dès le début de sa biographie. Bien que le changement de nationalité pour les personnes étrangères soit possible, il est encore plus fréquent que des double-nationaux (avec deux nationalités étrangères) soient inscrits dans les registres administratifs reçus comme source pour la production STATPOP, avec l'une ou l'autre nationalité. Par conséquent, il est possible que la nationalité change d'une année à l'autre dans STATPOP, selon la source dans laquelle est pris l'enregistrement. Dans les cas de multiples nationalités étrangères, c'est la nationalité présente dans les données les plus récentes qui sera prise en compte.

Si une personne est suisse dans sa dernière présence dans les données STATPOP et qu'elle est aussi suisse dans les données les plus anciennes, seule la nationalité suisse est prise en compte pour sa

biographie, quelles que soient les nationalités indiquées dans les présences intermédiaires dans STATPOP pour cette personne. Si une personne est suisse dans sa dernière présence dans les données STATPOP et qu'elle est étrangère dans les données les plus anciennes, la transition entre la nationalité étrangère et suisse est prise en compte lors de la première apparition de la nationalité suisse.

Pour toutes les personnes qui ont un changement de nationalité autorisé selon nos règles (d'étranger à Suisse), un mouvement d'acquisition de la nationalité suisse est recherché dans les données harmonisées. Si un tel mouvement existe, il est repris tel quel dans la statistique longitudinale. Si aucun mouvement d'acquisition n'est trouvé pour la personne dans les données STATPOP, un mouvement d'acquisition est imputé. La date d'acquisition correspondra alors au jour précédent le jour de référence des premières données dans lesquelles la nationalité suisse a été présente (par exemple si la personne est étrangère depuis 2010 et qu'elle est suisse pour la première fois dans les données du 30 septembre 2016, la date du mouvement d'acquisition sera imputée au 29.09.2016).

Les acquisitions qui ont lieu entre une émigration et une immigration ne sont pas gardées comme un mouvement, car seul les mouvements qui ont lieu lorsque la personne est domiciliée en Suisse sont considérés. Cependant, si une personne se naturalise à l'étranger, elle aura la nationalité suisse dans l'immigration qui suivra l'acquisition.

6.2.3 Traitements des autorisations de résidence

De manière générale, les autorisations de résidence pour étrangers sont reprises des effectifs trimestriels de STATPOP. S'il y a une différence d'autorisation de résidence entre deux effectifs successifs pour une personne, un mouvement de changement d'autorisation de résidence est créé avec la date (dateMove) du premier jour du trimestre durant lequel le changement d'autorisation de résidence a été observé. Si un changement d'autorisation de résidence est repéré durant un trimestre lors duquel une personne émigre puis immigré, le changement d'autorisation de résidence n'est pas imputé. L'autorisation de résidence de l'effectif du début du trimestre sera reprise dans l'émigration et l'autorisation de résidence de l'effectif de la fin du trimestre sera repris dans l'immigration.

Dans les immigrations et les naissances, l'autorisation de résidence correspond à celle du premier effectif trimestriel qui suit l'immigration ou la naissance dans lequel la personne est présente et dans les émigrations et les décès, l'autorisation de résidence correspond à celle du dernier effectif trimestriel dans lequel la personne est présente avant l'émigration ou le décès. Pour les immigrés sans présence dans un effectif trimestriel STATPOP en raison d'une présence de moins de trois mois en Suisse, l'autorisation de résidence dans les migrations est reprise des données des mouvements STATPOP si elle existe, dans le cas contraire, l'autorisation de résidence est sans indication (-9). Pour les naissances pour lesquelles aucune autorisation de résidence n'est disponible dans les effectifs trimestriels STATPOP en raison d'une émigration ou d'une acquisition de la nationalité suisse peu de temps après la naissance, l'autorisation de résidence est reprise de l'enregistrement de la mère ou de celui du père si le lien a pu être fait, dans le cas contraire, l'autorisation de résidence est sans indication.

Si une personne immigré et qu'elle n'a pas encore d'autorisation de résidence attribuée (code 13) dans sa première présence dans un effectif trimestriel, l'autorisation de résidence dans l'immigration correspond alors à celle du deuxième effectif suivant l'immigration, si celle-ci n'est pas attribuée, l'autorisation de résidence de l'immigration prend le code 13.

Il arrive parfois que les registres des habitants saisissent des codes 13 (pas attribué) dans la variable autorisation de résidence lors du renouvellement d'un livret. Les personnes qui ont une autorisation de résidence pas attribuée (code 13) dans un effectif trimestriel STATPOP, alors qu'elles ont eu des autorisations de résidence précédemment et ont une autorisation de résidence plus tard dans leur biographie se voient supprimer le code 13 dans la statistique longitudinale. Si le code 13 est attribué dans

le tout dernier effectif (celui qui correspond à la date de référence de la statistique), l'autorisation de résidence est alors contrôlée dans l'effectif du 31 mars de l'année suivante. Si le code 13 n'apparaît plus dans l'effectif du 31 mars, le code 13 ne sera pas pris en compte pour la fin de la biographie dans la statistique longitudinale.

6.2.4 Traitement du type de population

Le type de population résidente permanente (populationType = 1) est attribué systématiquement à tous les mouvements des personnes suisses, des diplomates et fonctionnaires internationaux et des personnes avec un livret B, C, Ci ou L>12 mois. Le type de population résidente non permanente (populationType = 2) est attribué systématiquement à tous les mouvements de personnes étrangères sans indication d'autorisation de résidence (-9), aux personnes avec une autorisation de résidence pas attribuée (13) ou avec un livret L<12 mois (704).

Lors d'un changement d'autorisation de résidence qui implique un changement de type de population (p.ex. livret L<12 mois à livret B), le type de mouvement du changement d'autorisation de résidence (RZ) est modifié en mouvement combiné de changement d'autorisation de résidence et de type de population (RTZ). Ainsi les changements de type de population sont facilement identifiables.

Pour les mouvements des personnes du domaine de l'asile (livrets N et F), la durée de la présence en Suisse détermine le type de population. Suite à son immigration, la personne se trouve dans la population résidente non permanente durant les 12 premiers mois de présence en Suisse. Un an après l'immigration, un mouvement de changement de type de population (moveType = TZ) est créé et la personne est alors considérée dans la population résidente permanente.

La règle des 12 mois de présence en Suisse est valable pour l'ensemble des mouvements des personnes du domaine de l'asile à l'exception des naissances. Pour les naissances avec un livret N ou F, le type de population est repris de l'enregistrement de la mère, si le lien a pu être fait, dans le cas contraire, le nouveau-né est considéré dans la population résidente non permanente durant les 12 premiers mois de sa vie.

6.3 Flags

Tous les enregistrements qui ont été créés ou imputés lors de la production des données longitudinales peuvent facilement être identifiés. À côté des cinq types de mouvements créés entièrement (naissances, décès, annulation des NAVS13, changements d'autorisation de résidence et changements de type de population), les enregistrements ajoutés dans l'effectif de 2010 ainsi que les mouvements imputés² sont marqués par une valeur dans le flag des enregistrements imputés (f_record).

Si la date de mouvement n'a pas pu être reprise d'une valeur existante dans un autre enregistrement de la personne concernée et qu'elle a été attribuée selon des algorithmes prédéfinis, cette imputation sera marquée dans le flag des dates de mouvement imputées (f_dateMove).

Tous les enregistrements qui sont repris de STATPOP et qui ne sont pas retenus dans les données finales, sont conservés dans une table séparées (remove_YYYYMMDD). Un flag (f_remove) permet d'identifier la raison pour laquelle l'enregistrement a été mis de côté (p.ex. doublons, immigration et émigration en moins de quatre jours, etc.).

² Enregistrements pour un type de mouvement donné qui sont créés en raison des incohérences dans les biographies et qui s'ajoutent aux enregistrements STATPOP repris de la KD.

7 Fichiers mis à disposition

Un fichier attributs_YYYYMMDD est disponible à la fin de la production des données longitudinales. Ce fichier contient d'une part les variables préparées dans la table des attributs de la phase d'harmonisation des données pour toutes les personnes présentes dans l'univers de la population considérée, et d'autre part de variables issues des biographies traitées, comme par exemple la date d'acquisition de la nationalité suisse (dateOfAcquisitionSwiss), la dernière date d'arrivée en Suisse (lastCHarrivalDate), la nationalité et l'autorisation de résidence dans le dernier enregistrement disponible dans la statistique longitudinale (lastNationalityState, lastResidentPermit), ainsi qu'un flag qui indique si la personne est présente en Suisse le jour de référence de la statistique (lastStock).

Un deuxième fichier event_YYYYMMDD contenant tous les mouvements (présences en 2010, naissances, décès, immigrations, émigrations, annulations des NAVS13, acquisitions de la nationalité suisse, changements d'autorisation de résidence et changements du type de population) des personnes présentes dans l'univers considéré est aussi disponible. Il contient les différentes variables pertinentes pour chacun des mouvements ainsi que des variables de qualité comme les flags et des variables construites facilitant les analyses (p.ex. âge au moment de l'événement, année du mouvement).

Les effectifs, recalculés pour chaque année au 31 décembre, ont aussi été préparés et sont disponibles (stockyyyy_YYYYMMDD, où yyyy correspond à l'année d'analyse des données et YYYY à l'année de référence de la statistique). Ces fichiers permettent de faciliter certaines analyses ou de calculer des indicateurs sans travaux de préparation de données préalables. Tous les enregistrements provenant des effectifs, des migrations ou des acquisitions STATPOP qui ne sont pas repris dans les données finales du module Nationalité se trouvent dans le fichier remove_YYYYMMDD.

8 Résultats

8.1 Résultats généraux

Dans le tableau 1 se trouve le nombre d'enregistrements considérés au total entre 2010 et 2020 : le nombre d'enregistrements par type de mouvements et le nombre d'enregistrements créés ou imputés dans la production de la statistique longitudinale.

Tableau 1 : nombre d'enregistrements par type de mouvement et selon leur origine

Type de mouvement	Nombre d'enregistrements par type		
	Nombre total	Enregistrements créés dans la DVS (non repris de STATPOP)	Taux créé dans la DVS (%)
Total	15 639 764	3 212 175	20,5
Présence en 2010	7 952 836	33 230	0,4
Immigrations	2 506 097	99 427	4,0
Emigrations	1 918 954	197 559	10,3
Naissances	857 290	857 290	100
Décès	662 394	662 394	100
Annulations du NAVS13	5 968	5 968	100
Acquisitions nationalité CH	388 035	8 117	2,1
Changements d'autorisation de résidence	1 260 994	1 260 994	100
Ch. de type de population	422 712	422 712	100

8.2 Valeurs manquantes

Le tableau 2 (ci-dessous) montre le taux de valeurs manquantes de différentes variables. On voit que, pour les variables démographiques, ce taux est très bas. Pour les variables date et pays de la dernière arrivée en Suisse, le taux élevé de valeurs manquantes s'explique par le fait que, pour les personnes nées en Suisse, il n'est pas possible de savoir si elles ont séjourné à l'étranger avant 2010, ces personnes ont par conséquent une valeur manquante alors que la très grande majorité sont en Suisse depuis leur naissance. Si on ne considérait que les étrangers qui ont immigré, le taux de valeur manquantes serait au-dessous du pourcent.

Pour les pays de provenance dans les immigrations et les pays de destination dans les émigrations le taux de valeurs manquantes est élevé. Ceci s'explique d'une part par le fait que dans les données STATPOP, les valeurs manquantes pour les pays de provenance et destination sont déjà élevées et d'autre part, par le nombre important d'immigrations et d'émigrations imputées pour lesquelles aucun lieu n'est disponible.

Tableau 2 : taux de valeurs manquantes par variable

Variable	Valeurs manquantes	Univers module Autorisation de résidence	Mouvements considérés	Taux de valeurs manquantes (%)
Date de naissance	0	10 666 287		0,00
Sexe	0	10 666 287		0,00
Pays de naissance	25 958	10 666 287		0,24
NAVS13 de la mère	5 362 268	10 666 287		50,27
NAVS13 du père	5 960 616	10 666 287		55,88
Commune de naissance pour les personnes nées en Suisse	17 818	6 735 858		0,26
Dernière date d'arrivée en Suisse	4 204 924	10 666 287		39,42
Pays de provenance lors de la dernière arrivée en Suisse	5 316 771	10 666 287		49,85
Dernière nationalité	3 372	10 666 287		0,03
Dernière autorisation de résidence pour les étrangers	17 551	3 430 950		0,51
Pays de provenance dans les immigrations	377 368		2 506 097	15,06
Pays de destination dans les émigrations	495 929		1 918 954	25,84

8.3 Comparaison des résultats avec STATPOP

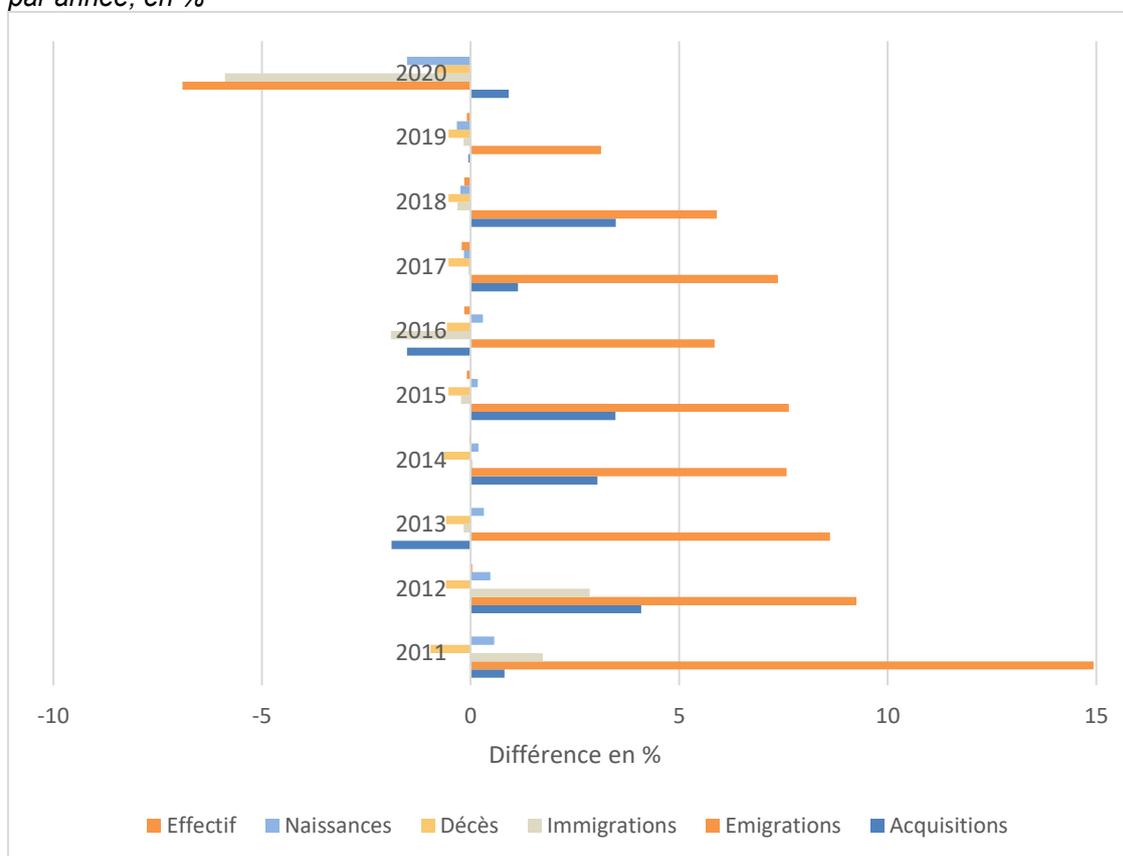
En comparant les résultats de STATPOP avec ceux de la statistique longitudinale, on remarque que les données sont très proches (voir graphique 1 et tableaux 3 à 8), à l'exception des émigrations. Les effectifs de fin d'année recalculés dans la statistique longitudinale ont au maximum 0,3% de différence avec les chiffres de STATPOP publiés.

Pour les émigrations, les différences sont plus marquées. En effet, pour les années 2011 à 2018, le nombre d'émigrations est en moyenne de 8,1% plus élevé que dans STATPOP. Cet écart peut s'expliquer par la méthode de relevé des données pour la production de STATPOP : les registres des habitants livrent trimestriellement à l'OFS les données des personnes présentes dans leur commune le jour de

référence ainsi que celles des personnes décédées ou parties de la commune lors des 12 mois précédant le jour de référence. Les annonces de départ de toutes les personnes parties à l'étranger sans avoir informé les autorités, pour lesquelles l'absence a été observée par le contrôle des habitants plus de 12 mois après le départ, ne seront jamais transmises à l'OFS et ne se retrouveront par conséquent pas dans les données STATPOP. Dans la statistique longitudinale, une émigration pourra être imputée pour ces cas.

Pour la dernière année disponible de la statistique longitudinale, on observe que les chiffres de cette dernière sont inférieurs à ceux de STATPOP pour l'ensemble des mouvements. Par exemple, cet état de fait pour 2019 a été modifié lorsque les données de l'année 2020 ont été ajoutées. En effet, dans STATPOP se trouvent les événements de l'année statistique ainsi que les événements de l'année précédente s'ils n'avaient pas été annoncés à temps. Dans la statistique longitudinale, tous les événements sont comptés dans l'année où le mouvement a eu lieu. Par conséquent, chaque année ajoutée complète particulièrement la dernière année avant l'année de référence (voir chapitre 8.4 pour plus de détails).

Graphique 1 : Différences entre STATPOP et statistique longitudinale pour l'effectif et les mouvements par année, en %



Dans les tableaux 3 à 8 se trouvent les résultats détaillés des comparaisons entre STATPOP (population résidente permanente et non permanente) et la statistique longitudinale.

Tableau 3 : comparaison des données STATPOP et statistique longitudinale pour l'effectif

	Effectif			
	STATPOP	DVS	Différence (N)	Différence (%)
2010	7 940 969	7 952 836	11 867	0,15
2011	8 035 391	8 034 991	-400	-0,00
2012	8 123 721	8 127 252	3 531	0,04
2013	8 222 327	8 222 627	300	0,00
2014	8 324 554	8 321 902	-2 652	-0,03
2015	8 423 612	8 416 116	-7 496	-0,09
2016	8 497 313	8 484 667	-12 646	-0,15
2017	8 560 151	8 542 133	-18 018	-0,21
2018	8 620 693	8 608 062	-12 631	-0,15
2019	8 681 911	8 673 674	-8 237	-0,09
2020	8 729 833	8 728 907	-926	-0,01

Tableau 4 : comparaison des données STATPOP et statistique longitudinale pour les naissances

	Naissances			
	STATPOP	DVS	Différence (N)	Différence (%)
2011	81 275	81 742	467	0,57
2012	82 844	83 235	391	0,47
2013	83 464	83 732	268	0,32
2014	85 747	85 909	162	0,19
2015	87 022	87 167	145	0,17
2016	88 383	88 648	265	0,30
2017	87 821	87 678	-143	-0,16
2018	88 215	88 003	-212	-0,24
2019	86 609	86 328	-281	-0,33
2020	86 135	84 848	-1 287	-1,52

Tableau 5 : comparaison des données STATPOP et statistique longitudinale pour les décès

	Décès			
	STATPOP	DVS	Différence (N)	Différence (%)
2011	62 201	61 609	-592	-0,96
2012	64 341	63 964	-377	-0,59
2013	65 024	64 652	-372	-0,58
2014	63 989	63 591	-398	-0,63
2015	67 675	67 318	-357	-0,53
2016	65 038	64 669	-369	-0,57
2017	67 032	66 676	-356	-0,53
2018	67 148	66 791	-357	-0,53
2019	67 853	67 492	-361	-0,53
2020	76 246	75 632	-614	-0,81

Tableau 6 : comparaison des données STATPOP et statistique longitudinale pour les immigrations

	Immigrations			
	STATPOP	DVS	Différence (N)	Différence (%)
2011	256 478	260 989	4 511	1,73
2012	262 175	269 884	7 709	2,86
2013	272 892	272 418	-474	-0,17
2014	262 695	262 838	143	0,05
2015	270 601	269 967	-634	-0,23
2016	249 812	245 138	-4 674	-1,91
2017	236 263	236 140	-123	-0,05
2018	245 490	244 724	-766	-0,31
2019	241 369	240 960	-409	-0,17
2020	214 997	203 039	-11 958	-5,89

Tableau 7 : comparaison des données STATPOP et statistique longitudinale pour les émigrations

	Emigrations			
	STATPOP	DVS	Différence (N)	Différence (%)
2011	168 776	198 398	29 622	14,93
2012	178 179	196 342	18 163	9,25
2013	178 715	195 573	16 858	8,62
2014	171 025	185 052	14 027	7,58
2015	180 169	195 058	14 889	7,63
2016	188 379	200 093	11 714	5,85
2017	183 533	198 125	14 592	7,37
2018	187 851	199 626	11 775	5,9
2019	187 835	193 900	6 065	3,13
2020	167 627	156 787	-10 840	-6,91

Tableau 8 : comparaison des données STATPOP et statistique longitudinale pour les acquisitions

	Acquisitions de la nationalité suisse			
	STATPOP	DVS	Différence (N)	Différence (%)
2011	36 012	36 307	295	0,81
2012	33 500	34 928	1 428	4,09
2013	34 061	33 430	-631	-1,89
2014	32 836	33 865	1 029	3,04
2015	40 689	42 153	1 464	3,47
2016	42 937	42 295	-642	-1,52
2017	44 949	45 469	520	1,14
2018	42 493	44 027	1 534	3,48
2019	41 127	41 107	-20	-0,05
2020	34 141	34 454	313	0,91

Les comparaisons des effectifs par année entre STATPOP et la statistique longitudinale ont aussi été réalisées par classe d'âge avec les données 2010-2018. Au total, la statistique longitudinale est en moyenne de 0,01% plus élevée que STATPOP. Pour toutes les années de référence, le nombre d'enfants est supérieur dans la statistique longitudinale. Pour les 0 à 18 ans, les chiffres de la statistique longitudinale sont en moyenne de 0,47% supérieurs à ceux de STATPOP. Si on ne considère que les 0 à 4 ans, le pourcentage atteint 0,71% ce qui correspond à 2 800 enfants de 0 à 4 ans par année en moyenne, entre 2010 et 2018.

Par contre, pour les classes d'âges supérieures, à l'exception de l'année de référence 2010, toutes les classes d'âges de toutes les années montrent un nombre de personnes supérieur dans STATPOP. La moyenne de la différence pour les années 2010 à 2018 pour les 18 ans et plus est de 0,17%. Elle atteint 0,27% si on ne tient compte que des personnes entre 18 à 44 ans.

8.4 Données rétroactives

Le fait que tous les éléments d'une biographie soient remis dans l'ordre chronologique, quelle que soit la date à laquelle l'information a été reçue, implique une mise à jour annuelle de toutes les données précédentes. Les données 2010-2018 ont ainsi été entièrement recalculées lors de l'introduction des données 2019 dans la production des données longitudinales. De manière générale, l'ajout des données d'une année supplémentaire engendre un nombre de modifications décroissantes en remontant dans les années.

Dans les tableaux 9 à 11, se trouve le nombre de modifications apportées par l'ajout des données 2019 dans les années 2010-2018 par rapport au module « Migrations » initialement calculé pour 2010-2018. Si on voit que, pour les effectifs, l'ajout d'une année supplémentaire n'a que très peu d'influence sur les années précédentes, les modifications sont bien plus significatives dans le cas des migrations. Pour les immigrations, l'augmentation est de 5% pour l'année précédente et pour les émigrations, l'augmentation est de 3,8% pour l'année 2017 et atteint 9% pour 2018.

Tableau 9 : différences dans les effectifs 2010-2018 avec l'ajout des données 2019

	Effectif			
	Sans 2019	Avec 2019	Différence (N)	Différence (%)
2010	7 952 346	7 952 603	257	0,00
2011	8 033 757	8 034 436	679	0,01
2012	8 125 769	8 126 620	851	0,01
2013	8 220 657	8 221 809	1 152	0,01
2014	8 319 236	8 320 822	1 586	0,02
2015	8 412 215	8 414 641	2 426	0,03
2016	8 478 158	8 482 451	4 293	0,05
2017	8 540 602	8 537 899	-2 703	-0,03
2018	8 618 437	8 610 956	-7 481	-0,09

Tableau 10 : différences dans les immigrations 2011-2018 avec l'ajout des données 2019

	Immigrations			
	Sans 2019	Avec 2019	Différence (N)	Différence (%)
2011	260 972	260 998	26	0,01
2012	270 041	269 972	-69	-0,03
2013	272 533	272 480	-53	-0,02
2014	262 941	262 891	-50	-0,02
2015	269 988	269 997	9	0,00
2016	245 062	245 152	90	0,04
2017	235 551	236 028	477	0,20
2018	232 013	244 235	12 222	5,00

Tableau 11 : différences dans les émigrations 2011-2018 avec l'ajout des données 2019

	Emigrations			
	Sans 2019	Avec 2019	Différence (N)	Différence (%)
2011	198 965	198 663	-302	-0,15
2012	196 637	196 467	-170	-0,09
2013	196 017	195 761	-256	-0,13
2014	185 702	185 310	-392	-0,21
2015	196 150	195 432	-718	-0,37
2016	202 460	200 773	-1 687	-0,84
2017	192 362	199 965	7 603	3,80
2018	174 129	191 910	17 781	9,27

Annexes

ANNEXE 1 : LIENS ENTRE LES MOUVEMENTS ET LEURS VARIABLES ASSOCIÉES

Le schéma A illustre chaque module thématique avec les mouvements (lignes) et les variables correspondantes (colonnes) prévus dans la statistique longitudinale.

Schéma A : Mouvements et variables envisagés dans la statistique longitudinale

	pseudoVN	stateOfBirth	municipalityIdOfBirth	sex	pseudoIdChild1-15	pseudoParents	lastPlaceOfOriginId1-5	dateOfBirth	CHArrivalDate	CHComesFromCountryId	departureDate	goesToState	dateOfDeath	dateAnnulVN	nationalityState	dateOfAcquisitionSwiss	typeAcquisitionSwiss	residencePermit	populationType	maritalStatus	dateOfMaritalStatus	cancelationPartnerReason	pseudoIdSpouse/partner	reportingMunicipalityId	comesFromMunicipalityId	arrivalDate	MovingDate	pseudoFederalBuildingId	federalDwellingId	geoCoordE/geoCoordN	PLZ	Flag_xx
Table attributs personnes	x	x	x	x	x	x	x	x																								
P10	x								x	x					x			x	x	x	x	x			x	x	x	x				x
Naissances	x							x							x			x	x	x	x	x			x	x	x	x				x
Immigrations	x							x	x						x			x	x	x	x	x			x	x	x	x				x
Décès	x												x					x														x
Emigrations	x										x	x						x														x
Annulation VN	x													x				x														x
Acquisitions CH	x														x	x		x														x
Changements d'autorisation de résidence	x																x	x														x
Changements état civil	x																	x	x	x	x	x										x
Changements commune	x																	x						x			x	x				x
Déménagements dans commune	x																	x					x			x	x	x				x
Table attributs bâtiments																							x				x			x	x	