

Statistiques sur le niveau d'instruction :

Comparaison de deux sources et tentative d'explication des différences

Introduction

En tant qu'institut de statistique, Statbel s'efforce de fournir une image représentative de la population en fonction de diverses caractéristiques sociodémographiques ou autres. L'une de ces caractéristiques est le niveau d'instruction de la population, mesuré par le diplôme le plus élevé obtenu. Pour ce faire, Statbel dispose de deux types de sources : d'une part, le recensement décennal, qui est entièrement réalisé à partir de sources administratives depuis 2011, et d'autre part, les différentes enquêtes menées par Statbel auprès des ménages belges et, en particulier, l'Enquête sur les forces de travail (EFT).

Ces deux types de sources tentent, chacune avec leur propre méthodologie, de donner la meilleure estimation possible du niveau d'instruction réel de la population belge. Cependant, les deux sources sont inévitablement confrontées à des imperfections et à des problèmes de qualité, qui sont de nature différente selon la source.

Ce document propose une analyse comparative des deux sources et tente d'expliquer les différences observées dans les résultats. Nous donnons d'abord un bref aperçu des principales caractéristiques des deux sources et de la manière dont elles tentent de mesurer le niveau d'instruction. Nous développons des explications possibles pour les différences observées. Enfin, nous réfléchissons à certaines considérations que les utilisateurs devraient prendre en compte afin de choisir la source de données adéquate pour leurs besoins spécifiques.

Source 1 : Recensement de la population (Census)

But et description sommaire

Le recensement de la population a pour objectif de dénombrer les logements et les résidents en Belgique à une date donnée, tout en recueillant des informations autour des thèmes suivants : les données démographiques, les ménages et les familles, l'instruction, l'emploi et les caractéristiques des logements et bâtiments.

Le Census 2021 répond au [Règlement européen d'exécution \(UE\) No 2017/543](#) de la Commission du 22 mars 2017 qui établit les règles pour l'application du [Règlement \(CE\) n°763/2008](#) du Parlement européen et du Conseil, concernant les spécifications techniques des thèmes des recensements de la population et du logement et de leur classification.

Méthodes de collecte des données, périodicité et population

Statbel, l'office belge de statistique, organise un recensement de la population tous les 10 ans. Depuis 2011, celui-ci s'établit entièrement sur base de données administratives et se nomme « Census ».

Le Census vise à récolter des données exhaustives sur toute la population.

Le niveau d'instruction le plus élevé est disponible dans deux types de sources : les sources administratives et les sources déclaratives¹. Les sources **administratives** comprennent les bases de données administratives relatives aux diplômes délivrés par les institutions compétentes en matière d'enseignement (Communautés et Régions), les données du dernier Censu (2011) et les déductions des diplômes faites sur base des inscriptions dans les institutions compétentes en matière d'enseignement (Communautés et Régions). Ces sources présentent l'immense avantage d'être hautement fiables. Lorsque des individus sont répertoriés dans ces sources, il est vraisemblable qu'ils aient réellement obtenu ce niveau de diplôme.

Les sources **déclaratives** rassemblent des bases de données dans lequel le niveau d'instruction déclaré par le répondant est disponible. Il s'agit du recensement de la population et des logements de 2001, de bases de données issues du marché du travail, des bases de données de la Banque Carrefour de la Sécurité Sociale pour les demandeurs d'emploi, des bulletins de naissance, des données de l'Enquête sur les forces de travail (EFT), de l'enquête sur les revenus et conditions de vie (SILC) et de l'enquête Adult Education Survey (AES).

Le Censu 2021 étant un recensement administratif, la priorité a été donnée aux données des bases de données administratives. Les informations des bases de données déclaratives n'ont donc été prises en compte que lorsque le niveau d'instruction pour une personne n'était pas disponible dans les bases de données administratives². Pour la population des 15 ans et plus en 2021, 70% des niveaux d'instruction sont issus du Censu 2011, 25% des autres sources administratives et 5% des sources déclaratives.

Définition et classification

Le niveau d'instruction désigne le niveau d'études le plus élevé atteint avec succès dans le système éducatif du pays où l'instruction a été reçue. Toute instruction formelle qui contribue à ce niveau doit être prise en compte même si elle a été reçue en dehors d'écoles et d'universités (comme les Centrum voor Basiseducatie, Centrum voor Deeltijds Onderwijs, Centrum voor Volwassenenonderwijs (CVO)).

La nomenclature utilisée se base sur la classification en vigueur le 1^{er} janvier 2021³ :

CITE niveau 0 : Inférieur à l'enseignement primaire

CITE niveau 1 : Enseignement primaire

CITE niveau 2 : Enseignement secondaire inférieur

CITE niveau 3 : Enseignement secondaire supérieur

CITE niveau 4 : Enseignement post-secondaire non supérieur

CITE niveau 5 : Enseignement supérieur de cycle court

CITE niveau 6 : Baccalauréat ou équivalent

CITE niveau 7 : Master ou équivalent

CITE niveau 8 : Doctorat ou équivalent

¹ Cf. Annexe 1

²

³ https://statbel.fgov.be/sites/default/files/files/documents/Census2021/EDU_METADATA_CENSUS_2021_FR.pdf Cf.

<https://statbel.fgov.be/fr/propos-de-statbel/methodologie/classifications/classification-internationale-type-de-education>

Les personnes âgées de 15 ans et plus sont classées dans une seule des catégories entre le niveau CITE 0 et CITE 8, en fonction du niveau d'instruction qu'elles ont atteint (le niveau le plus élevé atteint). La classification dans ces catégories se fait sur la base des cartographies de la CITE qui sont définies par les communautés. Les personnes de moins de 15 ans sont classées dans la rubrique "Sans objet"

La catégorie CITE 0 vise les personnes n'ayant pas terminé avec succès le niveau 1 de la CITE. Sont classées dans cette catégorie les personnes qui n'ont jamais suivi un programme éducatif, ont suivi un enseignement durant leur petite enfance (défini comme le niveau 0 de la CITE dans la classification des programmes éducatifs) ou ont suivi un enseignement primaire mais sans terminer avec succès le niveau 1 de la CITE.

En plus de cette classification détaillée, une division en trois classes est également souvent utilisée (faible, moyen et élevé). Les personnes ayant un niveau d'instruction « faible » sont les personnes qui détiennent au mieux un diplôme de l'enseignement secondaire inférieur. Les personnes ayant un niveau d'instruction « moyen » ont obtenu un diplôme de l'enseignement secondaire supérieur mais pas de diplôme de l'enseignement supérieur. Les personnes ayant un niveau d'instruction « élevé » ont un diplôme de l'enseignement supérieur.

Données manquantes

On peut distinguer deux types de données manquantes : d'une part, les bases de données que Statbel n'a pas (encore) pu obtenir et, d'autre part, les diplômes obtenus à l'étranger.

Les bases de données manquantes

Malgré le grand nombre de sources administratives que Statbel a pu obtenir⁴, un certain nombre de bases de données des institutions délivrant des diplômes manquent encore. Il s'agit :

- o pour les diplômes de l'enseignement primaire
 - o des jurys en Fédération Wallonie-Bruxelles
 - o de la Communauté flamande
 - o des écoles privées
 - o des écoles européennes
- o pour les diplômes de l'enseignement secondaire
 - o de l'enseignement de promotion sociale en FWB
 - o du premier et second degré de l'enseignement secondaire en FWB
 - o du troisième degré de l'enseignement des apprentis et chefs d'entreprise de la Région Bruxelles-Capitale (efp & sfpme) et de la Communauté germanophone à partir de 2017/2018
 - o des écoles privées
 - o des écoles européennes
- o pour les diplômes de l'enseignement supérieur
 - o de l'enseignement de promotion sociale en FWB
 - o des jurys en FWB
 - o des équivalences en FWB
 - o en Communauté germanophone à partir de 2018
 - o de l'enseignement des apprentis et chefs d'entreprise de la Région Bruxelles-Capitale (efp & sfpme) et de la Communauté germanophone à partir de 2017/2018
 - o des écoles privées.

⁴ Cf Annexe 1

Statbel continuera ses efforts et négociations afin d'obtenir ces bases de données manquantes pour le prochain Censu.

Les diplômés étrangers

Les diplômés obtenus à l'étranger ne sont présents dans les bases de données administratives que si une équivalence de diplôme a été accordée. Cela signifie que les valeurs manquantes sont nombreuses pour les personnes qui ont obtenu leur diplôme à l'étranger, soit parce qu'elles ont grandi dans un autre pays et n'ont émigré en Belgique qu'après leurs études, soit parce qu'elles ont obtenu leur plus haut diplôme dans un établissement d'enseignement à l'étranger.

Source 2 : Enquête sur les forces de travail (EFT)

But et description sommaire

L'enquête sur les forces de travail est une enquête trimestrielle auprès d'un échantillon de ménages privés représentatif de la population, menée pendant toute l'année. Elle repose sur les réponses de près de 110.000 personnes (répondants) de 15 à 89 ans. Son but principal est de classer la population de 15 à 89 ans en trois groupes (personnes occupées, chômeurs et inactifs), et de fournir des données descriptives et explicatives sur chacune de ces catégories.

Cette enquête est soumise à un règlement européen⁵. Eurostat, le service statistique de l'Union européenne, la coordonne auprès des différents États membres de l'Union Européenne. Le but est de collecter des informations comparables au niveau européen, notamment pour ce qui concerne les taux d'emploi et de chômage au sens des définitions du Bureau international du Travail (BIT), mais également de collecter et de diffuser des données non disponibles par ailleurs (par exemple la mobilité des travailleurs, les raisons du travail à temps partiel, les différentes formes de travail temporaire, la profession) ou disponibles à une autre fréquence, telles que le niveau d'instruction de la population en âge de travailler.

Méthode de collecte des données et périodicité

Jusqu'en 2016, l'Enquête belge sur les forces de travail était une enquête transversale où chaque répondant n'était interrogé qu'une seule fois lors d'un entretien en face à face avec un enquêteur muni d'un ordinateur (CAPI)⁶. Exceptionnellement, à savoir lorsque le ménage n'était composé que de personnes de 65 ans ou plus qui ne travaillaient plus, une enquête par téléphone (CATI) était également autorisée. Les outils utilisés étaient alors les mêmes que ceux de l'enquête CAPI.

En 2017, l'EFT belge est devenue une enquête par panel avec un modèle 2(2)2 dans lequel les répondants sont interrogés pendant deux trimestres consécutifs, ne le sont pas pendant les deux trimestres suivants et sont enfin à nouveau interrogés pendant deux trimestres consécutifs. Une collecte de données en mode mixte a été introduite en 2017, où les répondants sont interrogés principalement en face à face via CAPI lors de la première vague (= première interrogation), et par internet (CAWI) ou téléphone (CATI) lors des vagues suivantes. L'enquête peut être réalisée en quatre

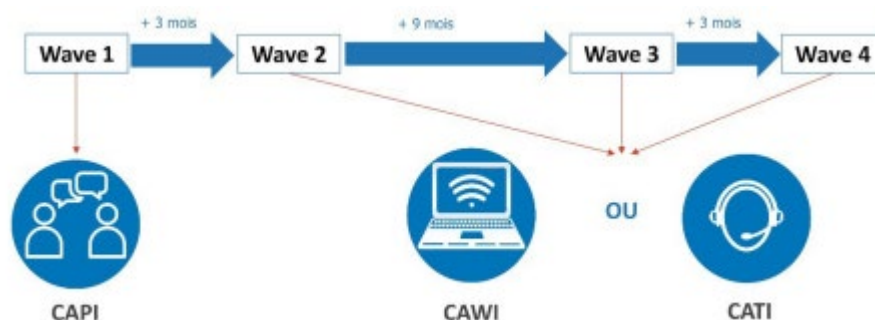
⁵À partir du 1^{er} janvier 2021, l'enquête sur les forces de travail a été adaptée au nouveau règlement-cadre du règlement (UE) [2019/1700](#). L'enquête est coordonnée par Eurostat, l'office statistique de l'Union européenne. Comme dans de nombreux autres États membres de l'UE, l'EFT en Belgique est une enquête obligatoire régie par l'arrêté royal du 10 janvier 1999 relatif à une enquête par sondage sur les forces de travail, modifié par l'arrêté royal du 25 mars 2016.

⁶ CAPI= Computer Assisted Personal Interviewing, CATI= Computer Assisted Telephone Interviewing, CAWI=Computer Assisted Web Interviewing

langues : le néerlandais, le français, l'allemand et l'anglais, afin de minimiser la barrière linguistique pour la participation.

Le processus de collecte de données est schématisé ci-dessous :

Schéma 1 : Collecte de données en mode mixte avec un panel 2(2)2.



Définition et classification

Le niveau d'instruction fait référence au niveau d'enseignement le plus élevé que la personne déclare avoir achevé avec succès et pour lequel un diplôme, un certificat ou une attestation a été délivré par les autorités compétentes en matière d'enseignement.

L'EFT mesure le diplôme le plus élevé obtenu au moyen d'un questionnaire complet. Lors de la première interrogation, les personnes interrogées doivent d'abord renseigner à l'enquêteur si elles ont suivi un enseignement primaire, secondaire ou supérieur pendant le mois de référence. Si le répondant a des doutes, l'enquêteur détermine avec lui le niveau approprié. Ensuite, différentes sous-questions permettent de déterminer le niveau d'enseignement le plus élevé terminé avec succès, le type et l'année.

Le questionnaire est disponible à l'adresse suivante: <https://statbel.fgov.be/fr/themes/emploi-formation/marche-du-travail/emploi-et-chomage#documents>

Les **certificats ou diplômes obtenus à l'étranger** constituent un cas spécifique. Ces diplômes doivent être classés selon le système d'enseignement belge. En cas d'absence de document d'équivalence pour le certificat/diplôme en question ou s'il est difficile d'attribuer un niveau, l'enquêteur peut se baser sur l'âge auquel le répondant a obtenu le niveau d'enseignement le plus élevé à l'étranger. Pour faciliter le classement des certificats ou des diplômes obtenus à l'étranger selon le système d'enseignement, l'enquêteur se base alors sur une table avec des âges théoriques d'obtention des différents diplômes.

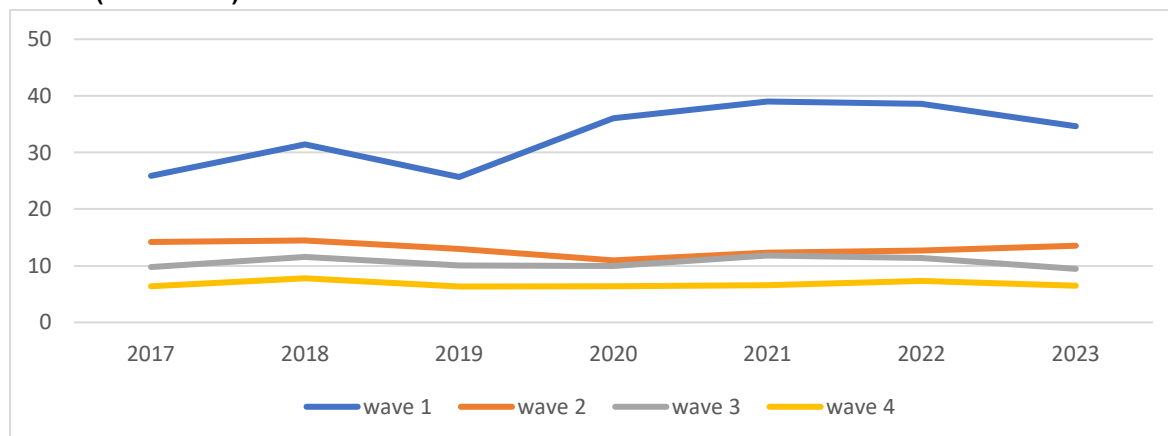
Non-réponse

Comme la plupart des enquêtes auprès des ménages ou des citoyens, l'EFT est confrontée à de la non-réponse. Cette situation se produit pour diverses raisons. Contrairement à d'autres enquêtes et, grâce à son caractère obligatoire, les refus ne sont pas la principale raison de la non-réponse, mais plutôt ce que l'on appelle les "non-contacts" (c'est-à-dire qu'aucun contact n'a pu être établi avec le ménage) ou des problèmes pratiques tels que l'abandon d'un enquêteur ou une pénurie d'enquêteurs dans certaines régions.

Le graphique ci-dessous montre l'évolution du taux de non-réponse par vague. Celui-ci est le plus élevé lors de la première vague de l'enquête. Dans la période précédant la pandémie de Covid-19, ce pourcentage se situait entre 26% et 31%, mais en 2020, le taux de non-réponse a augmenté pour

atteindre 36% en 2020 et 39% en 2021 et 2022. Ce n'est qu'en 2023 qu'une légère diminution a été observée. Lorsque les personnes ont participé à l'enquête une première fois, la probabilité d'une participation ultérieure est nettement plus élevée. Néanmoins, au cours de la deuxième vague, entre 11% et 14% des répondants abandonnent ; au cours de la troisième vague, ils sont en moyenne 11% et au cours de la quatrième vague, ils sont en moyenne 7%.

Graphique 1 : Enquête sur les forces de travail : Pourcentage de non-réponses, par vague, par année (2017-2023)



La participation ou non à une enquête est souvent liée à un certain nombre de caractéristiques sociodémographiques importantes, telles que l'âge, le niveau d'instruction, la situation au regard du logement, l'origine, etc. et peut donc fausser les résultats d'une enquête. Le calage (ou la pondération) de l'enquête vise à corriger autant que possible ce phénomène.

Brève explication des différences entre les résultats du Census et de l'EFT

Avec la publication des résultats du Census 2021, nous disposons à Statbel de deux sources qui, indépendamment l'une de l'autre et sur la base de leur propre méthodologie, tentent de donner une image du niveau d'instruction de la population belge. Ce faisant, nous constatons qu'il existe certaines différences dans les résultats, ce qui n'est pas surprenant étant donné les méthodes très différentes utilisées pour collecter et traiter les données. Les deux sources visent à représenter la meilleure estimation possible de la réalité. Les deux sources ont leurs propres qualités, mais aussi, inévitablement, leurs défauts. Même si nous le souhaitons, nous devons partir du principe qu'aucune des deux sources ne fournit une représentation parfaite du niveau d'instruction réel de la population belge, mais plutôt la meilleure approximation possible dans un contexte donné.

Dans cette note, nous souhaitons nous attarder sur les principaux facteurs qui peuvent expliquer la différence entre les résultats. Il est important de réaliser que les différences que nous observons sont le résultat de nombreux effets différents, qui eux-mêmes ne vont pas nécessairement toujours dans une seule direction. Ce que nous observons dans les résultats n'est que la différence nette qui subsiste. Dans l'ensemble, nous voyons cinq explications principales de ces différences.

Deux explications sont associées au fait que les résultats de l'EFT sont basés sur une enquête par sondage. Premièrement, cela signifie que les résultats peuvent être soumis à des fluctuations d'échantillonnage. Cet effet est d'autant plus important que l'on se concentre sur des sous-populations plus petites, puisque l'échantillon considéré est également plus petit dans ce cas. Ces types de fluctuations sont plus susceptibles d'être dues au hasard et peuvent donner lieu à des surestimations et à des sous-estimations du niveau d'instruction réel.

Deuxièmement, une enquête par sondage comme l'EFT souffre également de non-réponse, ce qui peut entraîner un biais plus systématique dans les résultats et qui fait que le niveau d'instruction réel selon

les résultats de l'EFT est plus susceptible d'être surestimé. Nous avons pu clairement identifier cet effet en examinant le taux de réponse, en fonction du niveau d'instruction mesuré par le Censur (information dont nous disposons à la fois pour les répondants et les non-répondants). L'effet se manifeste dans chacune des trois régions et devient plus important à mesure que le taux de réponse à l'enquête diminue, un problème auquel l'enquête est davantage confrontée depuis la pandémie de Covid en 2020.

Cependant, même après avoir contrôlé l'effet de la non-réponse, nous constatons que des différences subsistent entre les mesures du Censur et celles de l'EFT. Pour cela, nous voyons trois explications possibles, dont les deux premières devraient être recherchées dans le Censur.

Un premier facteur (la troisième explication) est lié au fait qu'il n'a pas été possible de trouver dans les sources disponibles une valeur pour toutes les personnes de la population recensée. Globalement, cela représente 6 % de la population totale, mais il semble y avoir de grandes différences selon les régions, ce qui est inextricablement lié à la composition différente de la population des trois régions. Par exemple, la population bruxelloise est nettement plus jeune et la proportion de personnes d'origine étrangère y est plus élevée. La proportion de valeurs manquantes y est donc de 16%, contre 5% dans les deux autres régions. Nous avons pu observer que ces valeurs manquantes avaient un impact sur l'estimation du niveau d'instruction selon le Censur, dans la mesure où ces valeurs manquantes ne sont pas distribuées de manière aléatoire entre les différentes catégories de niveau d'instruction. En utilisant les informations sur l'éducation de l'EFT spécifiquement pour ces valeurs manquantes dans le Censur, nous avons pu montrer que l'impact de ces valeurs manquantes varie considérablement en fonction de la sous-population. Par exemple, les personnes d'origine étrangère pour lesquelles aucune valeur de Censur n'a pu être déterminée ont plus souvent un niveau d'instruction plus élevé, surtout si elles sont d'origine européenne. Dans la population des plus de 65 ans, les valeurs manquantes sont plus susceptibles d'avoir un niveau d'instruction inférieur à celui du groupe pour lequel une valeur a pu être déterminée.

Une deuxième caractéristique du Censur, qui nous fournit une quatrième explication de la différence avec les résultats de l'EFT, a également trait aux informations manquantes, mais cette fois pour les personnes auxquelles une valeur du Censur a pu être attribuée, mais qui peuvent cependant ne pas correspondre au diplôme le plus élevé obtenu par la personne, soit parce qu'une source particulière était manquante, soit parce que la personne a obtenu un diplôme à l'étranger. Cela crée également un biais systématique dans les résultats, à savoir que le niveau d'instruction réel est sous-estimé.

Enfin, une cinquième et dernière explication doit être recherchée dans les diverses erreurs de mesure, qui peuvent se produire tant dans l'enquête EFT que dans les sources déclaratives utilisées par le Censur. Il s'agit d'effets très divers, dans différentes directions, tels qu'une mauvaise compréhension des questions ou des instructions par le répondant et l'enquêteur, mais aussi des lacunes dans les catégories de questions et de réponses, dans lesquelles un grand nombre de générations différentes, diplômées dans des systèmes éducatifs souvent très différents, devraient être en mesure de se situer.

Quelle source choisir en fonction de quel objectif ?

Il est important que les utilisateurs des résultats du Censur et de l'EFT soient conscients de ces différents éléments qui peuvent influencer les résultats et expliquer les différences entre les deux. Cependant, les sources sont également complémentaires dans une certaine mesure et chacune a ses propres points forts. Pour aider l'utilisateur à choisir l'une ou l'autre source, nous résumons ci-dessous les principaux points forts.

L'Enquête sur les Forces de travail est l'enquête de référence pour les indicateurs européens. Si l'on veut comparer le taux de chômage belge ou le taux d'emploi par niveau d'éducation avec ceux d'autres pays européens, il est préférable d'utiliser l'EFT. L'EFT est une enquête trimestrielle qui utilise un échantillon représentatif de la population. Elle permet de tracer des évolutions annuelles de 1986 jusqu'à maintenant. Les données annuelles de l'EFT sont disponibles rapidement après la fin de l'année

de référence (en mars) au niveau régional et provincial. Pour obtenir des croisements entre le niveau d'instruction d'une part, et les professions, le pourcentage de temps de travail, l'ancienneté dans un travail et encore beaucoup d'autres variables relatives au marché du travail d'autre part, il est nécessaire d'utiliser l'EFT.

Le Census est mené tous les dix ans et vise à recueillir des données exhaustives sur toute la population. Il est particulièrement intéressant pour les niveaux géographiques fins. Il est le seul à permettre de réaliser des descriptions sociodémographiques de la population au niveau des communes, des secteurs statistiques, pour comparer les niveaux d'instruction, selon les tranches d'âge, les situations familiales, etc. Les données sont disponibles tous les dix ans (un peu plus de trois ans après la date de référence). Le Census permet également d'analyser les comportements des individus en matière d'éducation au regard de leurs caractéristiques : relation éducation-emploi, relation éducation-conditions de logement, etc. Le Census sera donc utilisé lorsque des données au niveau détaillé sont nécessaires ou lorsque l'objectif est de croiser le niveau d'instruction avec des informations administratives.

Annexe 1 : Aperçu des bases de données obtenues, traitées et intégrées pour déterminer le **plus haut niveau d'instruction** obtenu dans le cadre du Censu 2021 (2011 → 2021)

Database	09-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	17-18	18-19	19-20	20-21	21-22	22-23	Description
Fédération Wallonie-Bruxelles															
COMPTAGES															Diplômes de l'enseignement obligatoire (primaire et secondaire) Diplômes précédents donnant accès à une inscription dans l'année supérieure
JURY ⁷															Diplômes obtenus via le jury central
SATURN															Diplômes de l'enseignement supérieur non universitaire Diplômes précédents donnant accès à une inscription dans le supérieur
CREF															Diplômes de l'enseignement supérieur universitaire Diplômes précédents donnant accès à une inscription dans le supérieur
EQUIVALENCES EO															Equivalences accordées dans l'enseignement obligatoire (primaire et secondaire)
Communauté germanophone															
DGOV															Diplômes de primaire, secondaire et supérieur (non-universitaire) ⁸
DGOV_ZAWN															Diplômes en apprentissage et chef d'entreprise
Région wallonne															
IFAPME															Diplômes en apprentissage et chef d'entreprise
Communauté flamande															
LED															Diplômes délivrés dans l'enseignement néerlandophone tout niveau confondu Diplômes déclarés par les demandeurs d'emploi en Flandre Equivalences du secondaire et du supérieur depuis 2018

⁷ Pour le jury en Fédération Wallonie-Bruxelles, Statbel a obtenu des données pour la période 2001-2016 puis de 2018-2022. Plus précisément, il s'agit de fichiers séparés pour chaque type de certificat (diplôme du premier cycle de l'enseignement secondaire, CE2D, CESS d'orientation générale et CESS de qualification technique et professionnelle)

⁸ Les diplômes de Hautes écoles (BAC) et Soins infirmiers (EPSC) n'ont pas été livrés pour les années scolaires 2005-2006 à 2008-2009 et de 2018-2019 à 2021-2022.

NARIC ⁹																Equivalences entre 2014 et 2018 pour les diplômes de l'enseignement secondaire et supérieur. Repris dans le LED après 2018
AUTRES – SOURCES ADMINISTRATIVES																
Ecole Royale Militaire																Diplômes délivrés par l'Ecole Royale Militaire Autre diplôme de l'enseignement supérieur obtenu précédemment
Census 2011																Plus haut niveau de diplôme atteint avec succès par les résidents en Belgique au 01/01/2011.
DEDUCTIONS																Déduction de diplômes sur base des inscriptions issues de la base de données des inscriptions
		2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022		
AUTRES – SOURCES DECLARATIVES																
Recensement 2001																Plus haut niveau de diplôme atteint avec succès par les résidents en Belgique au 01/01/2001
Banque Carrefour de la Sécurité Sociale																Données concernant les demandeurs d'emploi (FOREM, ADG, VDAB & ACTIRIS)
Base de données du marché du travail																Données collectées auprès d'entreprises au sujet de leurs travailleurs et données d'enquêtes (AES et SILC)
Bulletins de naissance																Niveau d'instruction du père/de la mère (issu d'eBirth)
Enquête sur les forces de travail (EFT)																Niveau d'instruction le plus élevé collecté dans le cadre de l'Enquête sur les forces de Travail

⁹ Statbel a obtenu toutes les équivalences reconnues entre 2014 et 2018, mais l'année des diplômes peut être de n'importe quelle année avant 2018. La base de données NARIC (= équivalences en Flandre) est intégrée entièrement dans le LED.