**Nomenclature des pays et des nationalités**

**dans l’enquête Trajectoires et Origines 2**

L’Insee et l’Ined ont assuré la co-maîtrise d’ouvrage de l’enquête Trajectoires et Origines pour sa première et sa deuxième édition (TeO1 2008-2009 et TeO2 2019-2020). Dans ce cadre, les deux instituts ont construit une **nomenclature des pays et des nationalités** pour permettre aux chercheurs et aux statisticiens publics de mener leurs travaux sur l’immigration en France. Cette classification vise à fournir des **regroupements pertinents du point de vue des vagues migratoires et des profils socio-culturels**. Ce faisant, sa mise en place et sa diffusion au-delà de l’enquête TeO2, doit permettre une meilleure **comparabilité** des résultats d’une étude à l’autre.

La nomenclature proposée est emboîtée **sur 4 niveaux[[1]](#footnote-2)** :

* **1 position** (6 modalités)**:** Ce niveau correspond aux continents, avec la distinction de la France.
* **2 positions** (10 modalités) : Ce niveau, très agrégé, fournit quelques subdivisions continentales. Il est utilisé principalement pour la diffusion de résultats couvrants toutes les origines. Si les données et les effectifs le permettent, il est possible de décomposer certains groupes en utilisant la classification au niveau 3, voire 4 pour distinguer certains pays en particulier.
* **3 positions** (27 modalités) : Ce niveau fournit le détail des principaux groupes de pays d’immigration. Ces regroupements ont été faits selon plusieurs critères : vagues migratoires, profils socio-culturels, effectifs, zones économiques etc. Ce niveau est utilisé pour des focus spécifiques à certaines origines ou, dans son ensemble, lorsque les tailles d’échantillons le permettent.
* **4 positions** (68 modalités) : Ce niveau fournit le détail des pays composant les groupes du niveau 3. Il vise à permettre d’effectuer des regroupements ad-hoc ou des monographies.

Une spécificité de TeO2 est de s’intéresser aux **personnes originaires des DROM-COM**, en les distinguant des personnes nées en métropole, car elles sont susceptibles d’être l’objet de discriminations similaires à celles des personnes issues de l’immigration. Cette distinction peut être également utile dans certaines analyses où le découpage administratif ne comprend pas l’ensemble du territoire français[[2]](#footnote-3). Toutefois, il n’est pas toujours utile ni souhaitable d’avoir cette distinction. Pour cette raison, il existe pour la France, une codification non-emboîtée distinguant la **« France, y compris DROM-COM » = 1000** et une autre qui distingue **la France métropolitaine et les territoires d’outre-mer (codes 1010 à 1030).**

****

**Comprendre et utiliser la nomenclature TeO2**

La **nomenclature « pays »** TeO2 est une **classification de territoires et non à proprement parler de pays.** Le statut juridique **des territoires d’outre-mer** de chaque pays détermine notamment sa classification. Par exemple, le Groenland, pays constitutif du Danemark, n’est pas classé comme ce dernier dans **410. Autres pays de l'UE 15.** En effet, s’il est considéré comme un territoire d'outre-mer associé à l'Union européenne, il n’en fait pas partie intégrante. Inversement, l’île de Madère (distinguée dans la nomenclature du TCM) est classée comme le Portugal au sein des pays de l’UE27. Pour les mêmes raisons, certains territoires d’Afrique ou d’Asie (au sens de la nomenclature internationale ISO), peuvent être classés dans l’UE et d’autres peuvent ne pas être classés avec leur pays de rattachement (par exemple, les Îles Vierges britanniques et les Îles Caïmanes, territoires d’Outre-Mer britanniques sont classées en Amérique du Nord).

La **nomenclature TeO2 des nationalités** diffère de celle des territoires dans sa construction bien que les codifications soient très proches. Par exemple : les Îles MALOUINES (territoire d’outre-mer britannique) sont classées dans :

* **5021 « Autres pays d'Amérique du Sud »,** dans la nomenclature TeO2 « pays »
* **4200 « Britanniques »,** dans la **n**omenclature TeO2 «nationalités »

**Pour mettre en œuvre** la nomenclature TeO2, des tables de passages, des catalogues de formats et des programmes de chargement sont mis à disposition des utilisateurs.

La construction peut se faire à partir de trois nomenclatures détaillées :

* **La codification ISO :** nomenclature internationale « ISO 3166 » sur 3 chiffres. Avec France = 250 et Italie=005. Elle existe également en 2 et 3 lettres (ISO-ALPHA).
	+ Il s’agit de la nomenclature utilisée dans le recensement et dans l’enquête Emploi depuis 2021. Des codes « spéciaux », commençant par « 9XX » sont parfois créés de manière ad-hoc. Par exemple, dans le recensement, le code « 999 »= « France, y compris DROM-COM » ou encore le code « 926 »=Kosovo dans l’EEC 2021, car ce dernier n’est pas reconnu par la nomenclature ISO mais l’est par Eurostat.
* **La codification du Tronc Commun des Ménages (TCM)** : nomenclature française et spécifique à l’Insee. Avec France = 001 et Italie=380. Contrairement à la nomenclature ISO, il existe des codes TCM différents pour certains pays et leur nationalité.
	+ Il s’agit de la nomenclature utilisée dans le Tronc Commun des Ménages des enquêtes ménages de l’Insee (TeO1 et TeO2, enquête Emploi jusqu’en 2020, SRCV etc.)
* **Le Code Officiel Géographique (COG)** : nomenclature française de pays et territoires étrangers, sur 5 caractères (commençant tous par « 99 » pour l’étranger). En théorie, la France n’y est pas codée mais le code « 99100 » est utilisée par convention. Italie = 99127
	+ Il s’agit de la nomenclature utilisée dans certains fichiers administratifs (Etat-civil par exemple) ou enquêtes anciennes.

**Six tables de passage** vers la nomenclature TeO2 sont fournies, pour 3 codifications possibles en entrée (ISO/TCM/COG) et 2 nomenclatures en sortie (pays/nationalités). Elles sont fournies sous Excel (*Correspondances\_format\_teo2.xlsx*) et SAS.

* **Attention**: les tables de passages ne sont pas « bijectives » d’une nomenclature à l’autre, notamment entre TCM et ISO. Bien qu’il existe de tables de passage de TCM vers ISO, certains territoires existent en codage TCM mais pas en ISO et inversement.

Sont également mis à disposition, des catalogues de format SAS (*teo2\_pays\_natio.sas7bcat*) et un programme (*3-Catalogue Formats TEO2.sas*) servant à les construire à partir des tables de passage SAS.

* Note : il est possible de créer d’autres formats de passage selon les besoins, par exemple avec en entrée des libellés pays/nationalités, à partir des tables de passage.



**Choix et pertinence des groupes d’analyse**

Par rapport aux classifications utilisées jusqu’ici, la nomenclature TeO2 introduit plusieurs nouveautés, notamment aux niveaux 2 et 3.

Tout d’abord, **l’Afrique** est subdivisée de la manière suivante au niveau 3 :



* **210. Afrique Sahélienne[[3]](#footnote-4)** et **211. Afrique centrale et du golfe de Guinée[[4]](#footnote-5)** : ces deux niveaux regroupent les principaux pays d’immigration d’Afrique subsaharienne. Les flux en provenance de ces pays se sont accélérés depuis la fin des années 1990. Toutefois, dans le cas de l’Afrique dite « Sahélienne », il s’agit d’une immigration moins qualifiée, plus masculine et plus endogame que celle originaire d’Afrique centrale et guinéenne. L’Afrique Sahélienne est composée pour une grande partie des pays de l’ancienne Afrique Occidentale Française (AOF). Alors que l’Afrique centrale et du golfe de Guinée regroupe des pays de l’ancienne de l’Afrique Équatoriale Française (AEF) et de pays du golfe de Guinée qui n’étaient pas forcément d’anciennes colonies françaises.
* **212.** **Comores, Madagascar :** regroupe deux pays dont l’immigration est numériquement importante en France.

Pour **l’Asie**, la classification au niveau 3 est la suivante :

L’une des principales nouveautés est la création du groupe « **301. Moyen-Orient**» dont l’immigration a augmenté ces dernières années en raison de la situation géopolitique. Il n’existe pas de définition arrêtée et consensuelle des pays composant le Moyen-Orient[[5]](#footnote-6). Dans la nomenclature TeO2, la **300. Turquie** est isolée du Moyen-Orient en raison de son importance numérique et de son histoire migratoire plus ancienne en France. Ces deux groupes sont toutefois réunies au niveau 2 (**30. Turquie, Moyen-Orient**), ce qui permet de les distinguer des autres pays asiatiques. À noter que **Chypre**, parfois inclus dans le Moyen-Orient en est ici exclu, car faisant partie de l’UE27. De même, le Moyen-Orient n’est ici que composé de sa partie asiatique. **L’Égypte**, pays d’Afrique, en est exclu. Toutefois, pour les chercheurs qui souhaiteraient reconstituer des groupes du type **MENA/WANA (Middle East and North Africa / West Asia and North Africa)**, comme cela est utilisé dans les études internationales, il est possible d’effectuer des regroupements ad hoc en utilisant les pays du niveau 4 (**2130. Libye, Égypte et 20. Maghreb**).

La classification distingue enfin au niveau 3 trois groupes de pays : **310. L’Asie du Sud-Est** dont l’immigration fut importante dans les années 1970/1980 ; **311. La Chine** dont l’immigration est plus féminisée et récente ; et enfin **312. L’Asie du Sud[[6]](#footnote-7)** dont l’immigration est également récente et importante en nombre.

**L’Arménie, l’Azerbaïdjan et la Géorgie** sont classées comme des pays d’Asie dans les nomenclatures internationales (ISO). Cependant, du fait qu’ils soient membres du conseil de l’Europe, ils sont classés dans la catégorie «**4211. Autres pays d’Europe** ».

**Effectifs dans le recensement**

Les tableaux suivants donnent les effectifs d’immigrés issus du recensement selon les différents niveaux de la nomenclature TeO2.

* Note : Depuis 2022, les chiffres-clefs sur le site de l’Insee utilise la nomenclature TeO2 de niveau 3. Ces séries sont actualisées chaque année.

Tableau : Distribution des immigrés par lieu de naissance, niveau 1 et 2 (en milliers)



Tableau : Distribution des immigrés par lieu de naissance, niveau 3(en milliers)



Tableau : Distribution des immigrés par lieu de naissance, niveau 4(en milliers)



1. On présente ici la classification des territoires et des régions. Celle sur les nationalités est identique mais tous les territoires français y sont regroupés dans « 1000 = Français » et il existe un code spécifique aux apatrides (6000). À noter que si les nomenclatures Teo2 territoires et nationalités sont quasi-identiques, les tables de passages pays et nationalités qui servent à les construire sont sensiblement différentes (cf. infa). [↑](#footnote-ref-2)
2. Pour rappel : France = France métropolitaine + 5 DROM. République française = France métropolitaine + 5 DROM + COM. Pour Eurostat, France = France métropolitaine + 5 DROM + Saint-Martin. [↑](#footnote-ref-3)
3. Sénégal, Mali, Guinée, Mauritanie, Burkina Faso,Tchad, Niger, Gambie, Guinée-Bissau [↑](#footnote-ref-4)
4. Côte d'Ivoire, République démocratique du Congo, Cameroun, Congo, Nigéria, Bénin, Ghana, Togo, République centrafricaine, Guinée équatoriale, Gabon [↑](#footnote-ref-5)
5. Dans la nomenclature TeO2, le Moyen-Orient est composé des pays suivants : Irak, Liban, Syrie, Iran, Israël, Jordanie, Arabie saoudite, Koweït,Bahreïn, Émirats arabes unis, Qatar, Oman, Yémen. [↑](#footnote-ref-6)
6. Inde, Pakistan, Sri Lanka, Maldives, Bangladesh, Népal, Bhoutan [↑](#footnote-ref-7)