

# Analyse statistique - Séance 2 : La catégorisation

Aubin Poissonnier-Beraud

Villes et environnements urbains - Université Lumière Lyon 2

2023

SIMON Patrick, « Les statistiques, les sciences sociales françaises et les rapports sociaux ethniques et de “race” », *Revue française de sociologie*, 2008, vol. 49, no 1, p. 153 162.

- ▶ Pourquoi créer des catégories statistiques sur la race ? Quels risques justifient l'opposition aux statistiques ethniques ou raciales ?
- ▶ Quelle voie a choisi la France ? Comment les chercheur-es travaillent sur ce rapport social ?
- ▶ Selon vous, les catégories statistiques doivent-elles s'inspirer des façons ordinaires dont les agents sociaux pensent et classent les individus ?

## Le fait migratoire

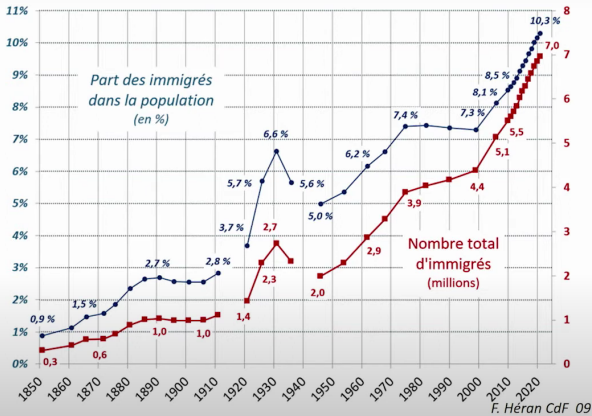
HERAN François, « Pour une vision historique et critique des droits des étrangers », 1er cours de la chaire *Migrations et sociétés*, année 2022-2023 : les migrations à la lumière du droit.

- ▶ Qu'est-ce que François Héran appelle le fait migratoire ?
- ▶ En quoi la fabrique du droit des étrangers est-elle liée à la mesure des flux démographiques ?

# L'immigration en France depuis 1850

Nombre et proportion d'immigrés en France depuis 1850 : une hausse générale, affectée par les aléas de l'histoire économique et juridique

Source :  
recensements  
SGF et Insee

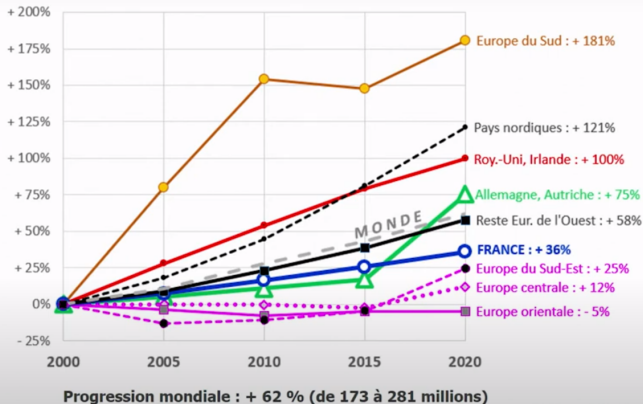


# L'immigration dans le monde depuis 2000

**Progression relative du nombre d'immigrés depuis 2000 dans le monde, en Europe et en France**

Source : ONU, *International Migrant Stock*, 2020

N.B. : définition ONU (inclut les nationaux nés à l'étranger)



F. Hérin CdF 11

## Compter : le bulletin individuel du recensement

Comment ces deux questions du recensement permettent-elles de compter le nombre d'immigré-es ?

**2 Date et lieu de naissance**

Né(e) le :   jour        mois          année

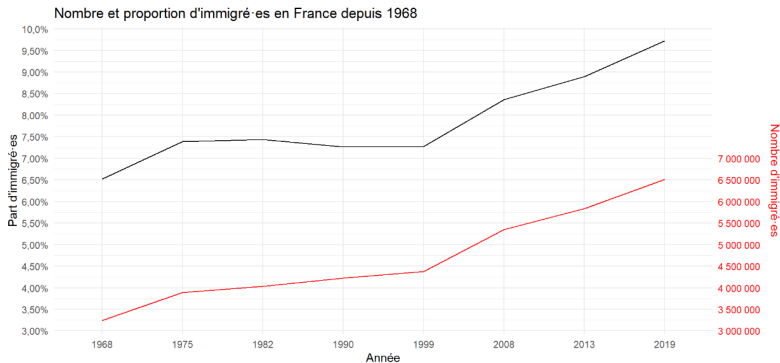
à :

commune (et arrondissement pour Paris, Lyon, Marseille)

département    n° DOM    pays pour l'étranger, territoire pour les COM

# Compter : des fichiers diffusés à l'analyse statistique

En imaginant la forme du tableau de données diffusé pour le recensement, quelles opérations de collecte, de recodage et de calculs ont abouti à ce graphique ?



Source : RP 1968-2019 (données harmonisées) | 2023 | A. Poissonnier  
Lecture : en 1968, il y avait 3 250 000 immigré-es en France et la part de proportion immigrée dans la population française était de 6,50%.

## Rappel de la définition

Appréhender un phénomène en le décomposant en des éléments plus simples, regroupés selon des logiques d'équivalence ou de similarité. C'est une opération de quantification qui part du langage pour construire des découpages du monde et dénombrer les individus qui tombent dans ces catégories.



## Le tri à plat pour compter à une dimension

Dans certains cas, le compte d'une catégorie constitue un objectif de l'analyse quantitative : connaître le nombre de personnes sans domicile fixe, mesurer la part de la population immigrée, déterminer les parts de diplômé-es du supérieur court et long.

Le *tri à plat* permet d'obtenir les effectifs de chaque modalité d'une variable catégorielle, c'est-à-dire le nombre d'individus présentant telle caractéristique (être diplômé-e du supérieur long) au lieu d'une ou plusieurs autres (être diplômé du supérieur court; ne pas être diplômé du supérieur).

## Décomposer la population active

Le tri à plat de la population active selon la variable catégorielle du statut d'emploi distingue les personnes en emploi, les personnes dans le halo du chômage et les personnes au chômage. Il permet de calculer le taux de chômage, statistique investie d'un enjeu politique fort.

Année	Population active	Au chômage	Taux de chômage	Dans le halo autour du chômage	Taux de halo du chômage	Taux total en mal-emploi
2022	30 575	2 234	7,3	1 859	6,1	13,4
2021 (r)	30 264	2 383	7,9	1 958	6,5	14,3
2020 (r)	29 735	2 392	8,0	2 241	7,5	15,6
2019 (r)	29 963	2 527	8,4	1 890	6,3	14,7
2018 (r)	29 977	2 705	9,0	1 878	6,3	15,3
2017	29 821	2 807	9,4	1 863	6,2	15,7
r : données révisées.						
Lecture : en 2022, 2 234 000 personnes sont au chômage et 1 859 000 dans le halo autour du chômage.						
Champ : France hors Mayotte, personnes actives de 15 ans ou plus vivant en logement ordinaire.						

Figure: Taux de chômage et de halo de chômage entre 2017 et 2022

## Le tri croisé pour comparer selon deux ou plus dimensions

Dans d'autres cas, la distribution d'une variable catégorielle est peu intéressante en elle-même mais elle le devient quand on la croise avec une ou plusieurs autres variables, soit quand on compare la distribution d'une variable considérée comme *expliquée* selon une ou plusieurs variables considérées comme *explicatives* : la part de femmes qui déclarent faire du bricolage comparée à la même part chez les hommes.

Le *tri croisé* permet d'obtenir les effectifs pour chaque couple de modalités de deux ou plus variables catégorielles, c'est-à-dire le nombre d'individus présentant telle caractéristique (être une femme) et telle caractéristique (bricoler) au lieu d'autres (être un homme et/ou ne pas bricoler).

# Étudier la composition sociale

	Agriculteurs, artisans, commerçants et chefs d'entreprise	Cadres et professions intellectuelles supérieures	Professions intermédiaires	Employés	Ouvriers	Retraités et inactifs	Ensemble <sup>(1)</sup>	Proportion de valeurs manquantes
Universités	8,8	33,1	15,3	17,6	10,3	14,8	100	7,9
dont : préparation DUT	9,8	30,4	18,4	20,4	12,1	9,0	100	3,1
formations d'ingénieurs <sup>(2)</sup>	8,5	45,1	16,2	13,3	7,9	9,0	100	3,6
autres formations	8,7	33,1	15,0	17,5	10,2	15,5	100	8,5
Formations d'ingénieurs hors université <sup>(2)</sup>	11,6	55,2	11,8	9,6	5,0	6,8	100	8,2
STS (scolaires)	11,7	15,5	14,6	19,6	22,5	16,0	100	9,7
CPGE et préparations intégrées	10,4	53,1	12,6	10,9	6,8	6,2	100	2,9
Écoles de commerce, gestion et comptabilité	18,6	49,4	8,7	11,6	4,8	6,9	100	37,7
Établissements privés d'enseignement universitaire	14,8	45,7	11,9	12,1	6,8	8,8	100	15,2
Écoles normales supérieures	6,8	62,6	10,6	7,5	2,1	10,3	100	5,3
Écoles artistiques, d'architecture et de journalisme	13,3	45,8	15,8	12,7	5,0	7,3	100	43,0
Écoles paramédicales et sociales <sup>(3)</sup>	15,1	23,5	9,3	31,3	18,7	2,0	100	12,0
Autres écoles et formations <sup>(4)</sup>	11,1	46,8	11,9	12,1	6,2	11,8	100	33,1
<b>Ensemble des étudiants français</b>	<b>10,5</b>	<b>34,7</b>	<b>14,1</b>	<b>17,2</b>	<b>10,9</b>	<b>12,6</b>	<b>100</b>	<b>13,1</b>

(1) Hors valeurs manquantes. (2) Y compris les formations d'ingénieurs en partenariat. (3) Les données concernant les écoles paramédicales et sociales sont celles de 2020-2021. (4) Comprend notamment les écoles juridiques et administratives.

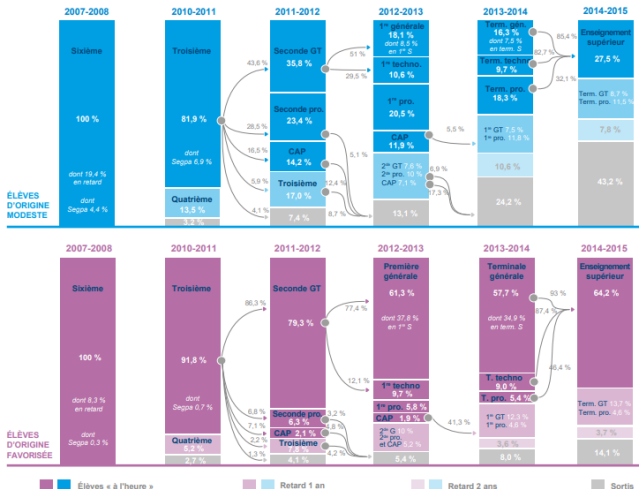
Champ : France métropolitaine et DROM.

Lecture : les étudiants dont les parents sont employés représentent 17,6 % des effectifs en universités en 2021-2022.

Source : MESRI-SIES, systèmes d'information SISE et SCOLARITÉ, enquêtes menées par le SIES sur les

# Étudier la différenciation sociale

Graphique 2 – Trajectoires comparées de la sixième à l'enseignement supérieur, selon l'origine sociale



- Notes :
- sont prises en compte ici les PCS ménage (nomenclature des professions et catégories socioprofessionnelles des ménages) ;
  - sont considérés d'origine modeste (35 % des élèves) les enfants de ménages à dominante ouvrière ou composés d'une personne ouvrière ou employée sans conjoint ou avec un conjoint inactif. Sont considérés d'origine favorisée (30 % des élèves) les enfants de ménages à dominante cadre ou intermédiaire/cadre ;
  - en sixième, sont dits en retard les élèves ayant redoublé au moins une année en primaire ;
  - pour les autres années, sont dits en retard d'un ou de deux ans les élèves ayant redoublé ou ayant été réorientés dans l'enseignement secondaire ;
  - Segpa : section d'enseignement général et professionnel adapté. Seconde GT : seconde générale et technologique.

Lecture : en 2010-2011, 81,9 % des enfants de famille modeste entrés en sixième en 2007 sont en troisième, 13,5 % sont encore en quatrième et 3,2 % sont sortis du système éducatif. Le faible écart entre cette somme et 100 % s'explique par l'existence d'élèves « perdus » dans le panel, décédés ou en très grand retard, ayant sauté une classe ou suivant une autre scolarité. Les flèches indiquent les taux de passage : 86,3 % des élèves de famille favorisée et 43,6 % des enfants de famille modeste qui étaient en troisième en 2010-2011 sont allés en seconde générale et technologique en 2011-2012. Pour simplifier, seuls quelques taux significatifs sont représentés.