



DATA  
INFRASTRUCTURE

# *Développer la culture des données*



MINISTÈRE  
DE L'ENSEIGNEMENT SUPÉRIEUR,  
DE LA RECHERCHE  
ET DE L'INNOVATION



Département données historiques, économiques et financières

Sandra Brée

# Rappels des missions

Mettre en place la politique des données quantitatives et issues d'enquêtes en sciences humaines et sociales en France

- **Produire**
  - Grandes enquêtes
  - Bases de données
- **Gérer**
  - Archiver
  - Diffuser
- **Promouvoir**
  - Faire connaître
  - accompagner

**PRO  
GE  
DO**

DATA  
INFRASTRUCTURE

# OUVRIR

# LA SCIENCE !

La science ouverte est la diffusion sans entrave des publications et des données de la recherche. Elle s'appuie sur l'opportunité que représente la mutation numérique pour développer l'accès ouvert aux publications et -autant que possible- aux données de la recherche.

<https://www.ouvrirelascience.fr/>

# OUVRIR

# LA SCIENCE !

Rendre les données

F  
indable

A  
ccessible

I  
nteroperable

R  
eusable



- Accompagnement sur la gestion des données (PGD - DMP)
- Sensibilisation aux questions liées à la protection des individus (RGDP)

- **Directeur** : Sébastien Oliveau
- **Animateurs des départements** :
  - Sandra Brée, "**Données historiques, économiques et financières**"
  - Frédéric Gonthier, "**Opinion et enquêtes sociopolitiques**"
  - Muriel Roger, "**Méthodes et statistiques publiques**"
  - Pascal Sébille, "**Population et parcours de vie**"
  - Philippe Terral, "**Santé**"
- **Archives issues de la statistique publique (ADISP)**: 4 Ingénieurs d'études
  - Lorraine Adam
  - Alena Lapatniova
  - Elodie Petorin
  - Ana Prada-Aranguren
  - Erik Zolotoukhine

# Gérer et diffuser les données

L'ADISP diffuse des enquêtes et bases de données produites par l'INSEE, plusieurs services statistiques ministériels et des institutions de recherche publiques. Les données diffusées sont accessibles gratuitement à l'ensemble de la communauté scientifique.

- Le catalogue compte aujourd'hui plus de 1.500 références.
- La DARES et la DREES veulent que l'ADISP soient diffuseurs pour leurs données auprès des organismes extérieurs.
- L'ADISP permet aussi d'obtenir des **produits sur mesure** (tableaux construits à la demande à partir d'une vingtaine de sources de l'INSEE (recensement, état civil, revenus, emploi, entreprises, etc.); donne accès aux **projections démographiques** de l'INSEE (Omphale) et au **distancier** METRIC.

## Données historiques, économiques et financières : Objectifs

- Permettre l'accès aux données historiques quantitatives (DOI)
- Soutenir les initiatives existantes ou en création pour permettre aux projets de mieux préserver et valoriser leurs travaux.
- Recueillir les bases de données => Adisp-Progedo
- Recenser les données existantes pour permettre leur identification et leur valorisation.

Création du département "Données historiques, économiques et financières" de Progedo

ARTICLE PRÉCÉDENT

Suivi du CoVid-19 en Méditerranée

CE CARNET VISE À PRÉSENTER LES ACTUALITÉS AUTOUR DE LA TGIR PROGEDO.

Rechercher...

Rechercher

ÉTIQUETTES

ADISP (8) Aix-Marseille (12) Base de données (2) Caen (5)  
Campus Condorcet (3) CDSP (9) colloque (2) Dijon (6) Données  
(2) Données historiques économiques et financières (3) ESS (5)  
EVS (3) Formation (3) GESIS (3) Grenoble (7) IDSC (2) INED (4)  
Interdisciplinarité (4) ISSP (3) Lille (8) Logiciel R (5) Logiciels  
libres (2) Lyon-Saint-Étienne (6) MSH (2) MSHD (2) Métadonnées  
(2) Nanterre (8) Nantes - Angers - Le Mans (14)  
Paris (6) Poitiers (3) PROGEDO (19) PROGEDO-  
Loire (10) PUD (34) PUDD (3) PUDL (3) PUDN (4)  
Quetelet-Progedo-Diffusion (7) Rennes (6) Santé (4)  
Science ouverte (2) Semaine Data-SHS (19) SHARE (3)  
Strasbourg (5) Toulouse (5) Université de Paris (2)

CATÉGORIES

Actualités (4)

# Les données historiques quantitatives disponibles en ligne

PAR SANDRA BRÉE · PUBLIÉ 27/04/2020 · MIS À JOUR 29/09/2021

## Les bases de données

- **sur Quetelet-Progedo-Diffusion :**
  - [Données sur la population et l'enseignement : la France au 19ème siècle](#)

**Pour le moment, une seule base de données historiques est disponible sur Quetelet-Progedo-Diffusion, mais d'autres sont en train d'être mises en ligne. Si vous avez des bases de données à proposer, écrivez-nous ([sandra.bree@cnrs.fr](mailto:sandra.bree@cnrs.fr))**
- **sur le site de l'INSEE sont disponibles :**
  - [Les Recensements de 1851 à 1921](#)
  - [Les Mouvement de la population, 1800-1925](#)
  - [Enseignement primaire, 1829-1897](#)
  - [Territoire et population, 1800-1890](#)
  - [Populations communales de 1876 à 2017](#)
- **sur l'Atelier du CRH sont disponibles :**
  - [Les recensements de 1901 à 1921](#)
  - [Les mouvements de la population de 1836 à 1925](#)
  - [L'enseignement primaire et secondaire de 1865 à 1906](#)
  - [La statistique industrielle de 1861 à 1896](#)
  - [L'enquête agricole de 1852](#)
  - [L'enquête postale de 1847](#)
- **SIG de la France de la Troisième République (TRF-GIS) :** <https://www.data.gouv.fr/fr/datasets/systeme-dinformation-geographique-de-la-france-de-la-troisieme-republique-1870-1940/>

Progedo

Quetelet Progedo Diffusion

CHAÎNE CANAL-U DE PROGEDO

2020 - Cartographier en 1h avec MAGRIT\* : Exemple avec les données du recensement - PUD-AMU

2020 - Traiter une base de données : l'essentiel pour débuter en statistiques sur SAS\* et ...

2020 - Présentation du CERGAM, analyse de séquences avec R et exemple d'utilisation de la Box ...

LES OUTILS QUANTI

Visualiser une analyse géométrique des données avec ggplot2 (R/RStudio)

Installer R et RStudio (presque !) sans peine

Tutoriel : Constituer un corpus Europresse utilisable dans R, Iramuteq et TXM

Tutoriel pour de l'analyse textuelle de réseau avec



- **sur le catalogue Nesstar, les données de nombreuses enquêtes de l'INED réalisées depuis les années 1950 sont en ligne :** <http://nesstar.ined.fr/webview/>  
L'INED a également rassemblé [une base de données rassemblant des séries longues](#) couvrant une cinquantaine d'indicateurs pour la France métropolitaine et la France entière.
- **Le site Didomena est l'entrepôt de données de la recherche de l'EHESS :**  
<https://didomena.ehess.fr/?locale=fr>
- **La base DFIH :** <https://dfih.fr/use>
- **La base Cassini :** [http://cassini.ehess.fr/cassini/fr/html/9\\_aide.htm](http://cassini.ehess.fr/cassini/fr/html/9_aide.htm)
- **L'enquête TRA :** <https://tra.site.ined.fr/>
- **Les tables de mortalité françaises pour les XIXe et XXe siècles :**  
[https://www.ined.fr/Xtradocs/cdrom\\_vallin\\_mesle/contenu.htm](https://www.ined.fr/Xtradocs/cdrom_vallin_mesle/contenu.htm) (source : Vallin et Meslé, 2001)
- **La French Human Mortality Database :** données sur la mortalité en France par région et départements depuis 1901 : <https://frdata.org/fr/french-human-mortality-database/>
- **La bibliothèque historique du Ministère de l'Agriculture (Bibagri)** rassemble un fonds patrimonial très riche constitué d'œuvres complémentaires, en termes de champs d'investigation, de nature de documents (administratifs, techniques, scientifiques...), voir ainsi :
  - [La statistique des préfets](#)
  - [La statistique de la France](#)
  - [Les enquêtes agricoles](#)
  - ainsi que des [Atlas agricoles numérisés](#)
  - et beaucoup d'autres [liens utiles](#)
- **De nombreuses données quantitatives sur le site de l'ICPSR, notamment :**
  - [Social, Demographic, and Educational Data for France, 1801-1897](#) (ICPSR 48)
  - [Violent Events in France, 1830-1860 and 1930-1960](#) (ICPSR 9080)
  - [Demographic, Social, Educational and Economic Data for France, 1833-1925](#) (ICPSR 7529)
  - [Prenuptial Agreements and Matrimonial Property Regimes in France, 1855-2010](#)
  - [Students in Secondary Schools in France, 1864](#)
  - [Analysis of Arrests in Paris, June 1848](#)

- [La statistique des préfets](#)
- [La statistique de la France](#)
- [Les enquêtes agricoles](#)
- ainsi que des [Atlas agricoles numérisés](#)
- et beaucoup d'autres [liens utiles](#)
- **De nombreuses données quantitatives sur le site de l'ICPSR, notamment :**
  - [Social, Demographic, and Educational Data for France, 1801-1897](#) (ICPSR 48)
  - [Violent Events in France, 1830-1860 and 1930-1960](#) (ICPSR 9080)
  - [Demographic, Social, Educational and Economic Data for France, 1833-1925](#) (ICPSR 7529)
  - [Prenuptial Agreements and Matrimonial Property Regimes in France, 1855-2010](#)
  - [Students in Secondary Schools in France, 1864](#)
  - [Analysis of Arrests in Paris, June 1848](#)
- **De nombreuses bases de données sur le site du LIR3S :**
  - Base de données du site *Académie/Académies* – [Accès au site](#)
  - Base de données de la répression judiciaire de la Commune de Paris- [Accès à la base](#)
  - Base de données des poursuivis à la suite du coup d'Etat de décembre 1851 – [Accès à la base](#)
  - Base de données des inculpés de l'insurrection de Juin 1848 – [Accès à la base](#)
  - Annuaire rétrospectif de la magistrature – XIXe-XXe siècles – [Accès à la base](#)
  - [CRIMINOCORPUS – Le portail sur l'histoire des crimes et des peines](#)
  - Base de données "Brochures du mouvement ouvrier" – [Accès à la base](#)
- L'AHCESR recense de nombreuses bases de données en ligne que vous pouvez retrouver ici : <https://ahcesr.hypotheses.org/bases-de-donnees>

# Les publications numérisées

## FRANCE

**Liste récapitulative de l'état civil numérisé** (archives nationales, archives nationales d'outre-mer, archives départementales et communales), sur le site de France Archives :

<https://francearchives.fr/fr/article/38170>

Données statistiques en open data de la statistique publique :

[https://francearchives.fr/fr/open\\_data](https://francearchives.fr/fr/open_data)

### Recensements de la population

en ligne sur Epsilon :

– [Dénombrements de la population de 1826 à 1896](#)

– [Dénombrements de la population de 1901 à 1936](#)

– [Recensement de la population de 1946](#)

– [Recensement de la population de 1954](#)

– [A partir de 1962, les recensements de la population sont disponibles sur l'Adisp](#)

### Annuaire statistique de la France

en ligne sur le site de Gallica :

– 1878-1899 (sauf 1893, 1894 et 1896) :

<https://gallica.bnf.fr/ark:/12148/cb343503965>

[/date.r=Annuaire%20statistique%20de%20la%20France](#)

– 1900-1940 ; 1946 ; 1951 :

<https://gallica.bnf.fr/ark:/12148/cb34350395t/date.r=Annuaire%20statistique>

– 1948 ; 1953-2007

<https://gallica.bnf.fr/ark:/12148/cb34349577d/date&rk=257512;0>

### Statistique annuelle du mouvement de la population

en ligne sur le site de Gallica (1893-1903 ; 1931) :

<https://gallica.bnf.fr/ark:/12148/cb42279947v/date>

### Statistique de la France : Territoire et Population

– [en ligne sur Epsilon : \(1801-1836\)](#)

### Les publications de la statistique publique

Plus de 40.000 documents produits par l'INSEE et ses ancêtres et partenaires depuis le 19e siècle.

### Mouvement économique en France

– [en ligne : 1932](#)

### Le Compte général de l'administration de la justice civile et commerciale

en ligne sur le site de Gallica : 1820-1940 (certains années manquantes) :

<https://gallica.bnf.fr/ark:/12148/cb345304266/date&rk=321890;0>

## PARIS ET LE DÉPARTEMENT DE LA SEINE

### Recherches statistiques sur la ville de Paris et le département de la Seine

(1821-1860)

A propos de ces publications : <https://bibliotheque.insee.net/statistiques-paris-comte-chabrol-de-volvic.aspx> en ligne, [Extraits des recherches statistiques sur la ville de Paris et le département de la Seine \(1824\)](#)

### Bulletin de statistique municipale de la ville de Paris

en ligne sur le site des bibliothèques spécialisées de la ville de Paris (1865-1879)

<https://bibliotheques-specialisees.paris.fr/ark:/73873/pf0000316261?>

### Annuaire statistiques de la ville de Paris

en ligne sur le site des bibliothèques spécialisées de la ville de Paris (1880-1967)

<https://bibliotheques-specialisees.paris.fr/ark:/73873/pf0000262800>

## AUTRES DÉPARTEMENTS


[L'enquête sur la population de Charleville du XVII<sup>e</sup> au XIX<sup>e</sup> siècle](#) a ouvert son site


Internet et notamment son portail des collections qui permet la consultation libre de milliers d'archives numérisées complémentaires de celles déjà accessibles sur le site des archives départementales, notamment celles sur lesquelles prend appui la base de données sociodémographiques. Un portail de données spatiales et démographiques sont en cours de construction.


# Les données économiques et financières

PAR **SANDRA BRÉE** · PUBLIÉ 16/10/2020 · MIS À JOUR 22/10/2020

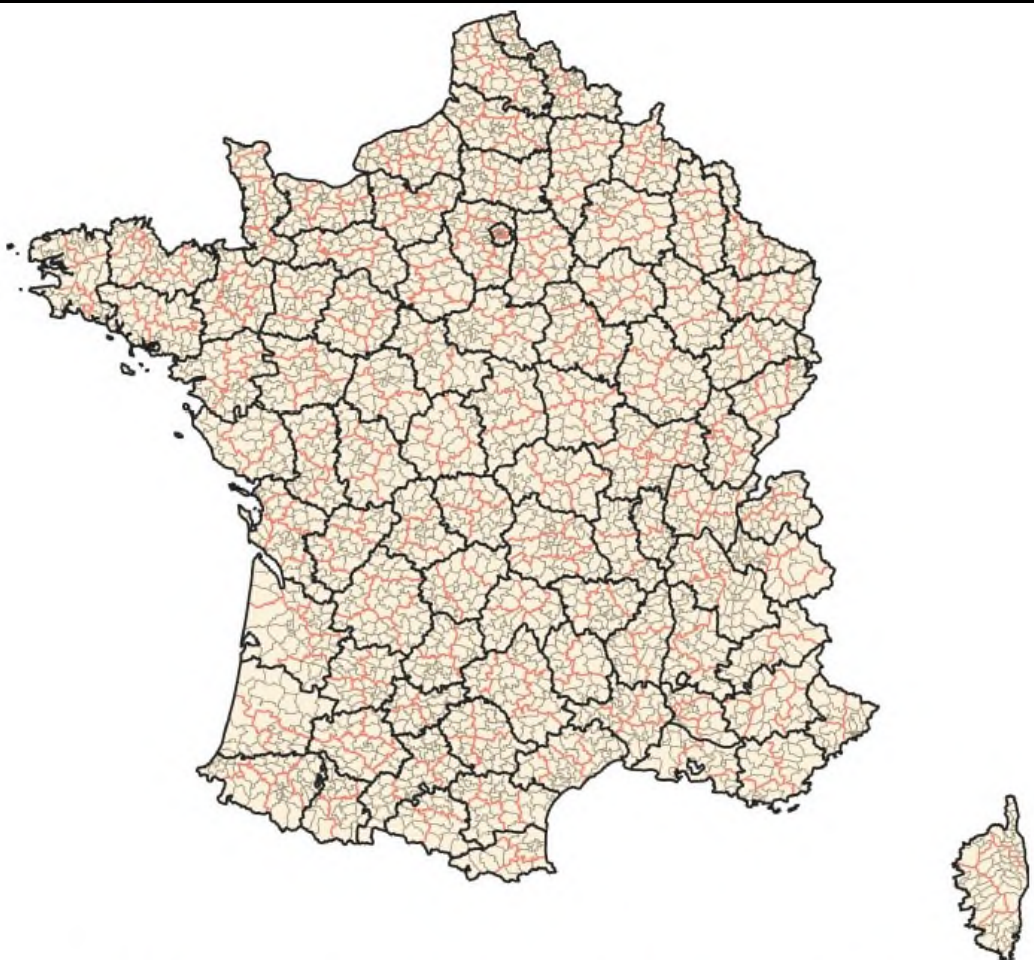
[DBnomics, the world's economics database](#) : 70 fournisseurs, 22 062 ensembles de données et 743 387 663 séries

[Eurofidai](#) : L'Institut Européen des données financières est un institut de recherche créé en 2003 par le CNRS (Centre National pour la Recherche Scientifique). Sa mission principale : développer des bases de données boursières européennes pour les chercheurs universitaires en finance.

[DFIH, Data for Financial History](#) : Une base de données complète sur les marchés boursiers français depuis 1795

[EurHisFirm](#) : vise à concevoir une infrastructure de recherche de classe mondiale pour collecter, fusionner, extraire, collationner, aligner et partager des données historiques détaillées de haute qualité au niveau des entreprises pour l'Europe. Pour atteindre cet objectif, il développe des outils innovants et déclenche la révolution des "grandes données" dans les sciences sociales historiques.

# SIG historique pour la France de la troisième République avec des nomenclatures annuelles et des shapefiles de 1870 à 1940 (Victor Gay)



(b) General administrative constituencies in 1871

La base de données TRF-GIS correspond à un système d'information géographique de la France de la Troisième République. Elle met à disposition nomenclatures et shapefiles annuels correspondant aux circonscriptions administratives de France métropolitaine de 1870 à 1940. Elle décrit les circonscriptions administratives générales (départements, arrondissements, cantons) ainsi que les circonscriptions militaires, judiciaires, pénitentiaires, électorales, académiques, ecclésiastiques, et les inspections du travail. Elle met aussi à disposition des nomenclatures annuelles établissant une correspondance entre chaque commune contemporaine et les circonscriptions auxquelles elle appartenait.



# Mariages parisiens à la fin du XIXe siècle : une micro-analyse quantitative (Maurice Garden)

DATA INFRASTRUCTURE

	A	B	C	D	E	F	G
1	NUMERO	DATEMARMAR	JOURMAR	MOISIRGMAF	PATROFEMME		PRENOMFEM1
2	1	18850103	03	01		GANDON	BERTHE
3	2	18850103	03	01		CURIE-NODIN	JEANNE
4	3	18850103	03	01		CHAPELAIN	JOSEPHINE
5	4	18850103	03	01		GUILLEMOT	SIMONE
6	5	18850103	03	01		GILSON	MARIE
7	6	18850103	03	01		LE BIGOT	MARIE
8	7	18850103	03	01		STEINER	HENRIETTE
9	8	18850105	05	01		LESPRIT	HONORINE
10	9	18850106	06	01		DE BAUDRIEUL	LOUISE
11	10	18850106	06	01		MIGNON	LOUISE
12	11	18850107	07	01		PASCOT DE LATOUCHE	MARIE
13	12	18850108	08	01		LANDEL	MARGUERITE
14	13	18850108	08	01		DAVID	EUGENIE
15	14	18850108	08	01		MARTIN	JULIENNE
16	15	18850108	08	01		POPPE	FANNY
17	16	18850108	08	01		DELEVOYE	BERTHE
18	17	18850108	08	01		SILVESTRE	MARIE
19	18	18850108	08	01		LANSADE	MARIE
20	19	18850110	10	01		CAYOLLE	MARIE
21	20	18580110	10	01		RICHET	MARIA
22	21	18850110	10	01		LEBRELL	EUGENIE
23	22	18850110	10	01		JEAN	JOSEPHINE
24	23	18850110	10	01		JACQUIN	EUGENIE
25	24	18850110	10	01		LALLEMAND	MARIE
26	25	18850110	10	01		DEL ROSARIO BLAY DE GOIG DE...	CLARA
27	26	18850110	10	01		TERRET	CECILE
28	27	18850110	10	01		ORTELLI	MARIE
29	28	18850110	10	01		PETIET	ATHENAIS

Le but de cette base construite à partir d'un échantillon de mariages parisiens de l'année 1885 était de « décrire la population présente dans ces actes, et essayer de comprendre comment se formait en partie la population parisienne en cette fin du XIXe siècle. » (Garden, 1998, p. 111)

1000 mariages ayant eu lieu pendant le mois de janvier 1886 dans les 20 arrondissements de Paris et 19 communes du département de la Seine : relevé des informations concernant les deux époux, leurs parents et les quatre témoins.

3 fichier de données :  
GMARI ; GFEMM ; GTEM

# Stature des conscrits français des classes 1800-1940 (Laurent Heyberger) Niveaux de vie et disponibilités alimentaires en France au milieu du XIXe s.

## 1. Stature des conscrits français des classes 1800-1940

1/ mesurer les évolutions et disparités de niveaux de nutrition nette des jeunes gens nés entre 1779 et 1920 à l'échelle individuelle à partir des listes de tirage au sort de la conscription ou leur équivalent;

2/ facteurs explicatifs : variables tirées de ces listes (profession et lieu de naissance -uniquement pour Mulhouse), des recensements publiés (milieu de vie rural/urbain) et de diverses variables qualitatives ou quantitatives collectées (salaires, productions agricoles, consommations);

La référence met à disposition 14 fichiers de données classées par département et par période.

## 2. Niveaux de vie et disponibilités alimentaires en France au milieu du XIXe siècle

1/ Indices classiques, du type salaires, disponibilités alimentaires ou budgets populaires pour expliquer ou non les disparités spatiales et sociales de stature (niveau de vie biologique).

2/ variables explicatives (données de l'enquête agricole de 1852, aux données du recensement et aux données anthropométriques individuelles des listes de tirage au sort) pour tester l'exception française -d'un point de vue anthropométrique- de grands citadins, dans un contexte occidental où l'urbanisation de l'industrialisation précoce se traduit plutôt par un malus urbain.

La référence met à disposition 1 fichier de données :  
**stature\_enq\_agri\_arrondissement** – qui compte 37 variables et 105324 observations

	A	B	C	D	E	F	G	H	I
1	départemé	ondissé	cantoi	professioni	csp	stature	spesvit	instructio	structionco
2	Alpes Hautes	Embrun	Chorges	boulangier	alim	60	31,2	12	oui
3	Alpes Hautes	Embrun	Chorges	charron	bois	73	31,2	12	oui
4	Alpes Hautes	Embrun	Chorges	cultivateur	agrindép	69	31,2	0	oui
5	Alpes Hautes	Embrun	Chorges	cultivateur	agrindép	32	31,2	12	oui
6	Alpes Hautes	Embrun	Chorges	cultivateur	agrindép	49	31,2	12	oui
7	Alpes Hautes	Embrun	Chorges	cultivateur	agrindép	55	31,2	12	oui
8	Alpes Hautes	Embrun	Chorges	cultivateur	agrindép	55	31,2	12	oui
9	Alpes Hautes	Embrun	Chorges	cultivateur	agrindép	61	31,2	12	oui
10	Alpes Hautes	Embrun	Chorges	cultivateur	agrindép	63	31,2	12	oui
11	Alpes Hautes	Embrun	Chorges	cultivateur	agrindép	64	31,2	12	oui
12	Alpes Hautes	Embrun	Chorges	cultivateur	agrindép	66	31,2	12	oui
13	Alpes Hautes	Embrun	Chorges	cultivateur	agrindép	69	31,2	12	oui
14	Alpes Hautes	Embrun	Chorges	cultivateur	agrindép	71	31,2	12	oui
15	Alpes Hautes	Embrun	Chorges	cultivateur	agrindép	75,5	31,2	12	oui
16	Alpes Hautes	Embrun	Chorges	cultivateur	agrindép	80	31,2	12	oui
17	Alpes Hautes	Embrun	Chorges	cultivateur	agrindép	80	31,2	12	oui
18	Alpes Hautes	Embrun	Embrun	boulangier	alim	64	31,2	12	oui
19	Alpes Hautes	Embrun	Embrun	charron	bois	57	31,2	12	oui
20	Alpes Hautes	Embrun	Embrun	charron	bois	66	31,2	12	oui
21	Alpes Hautes	Embrun	Embrun	clerc de notair	employé	64	31,2	12	oui
22	Alpes Hautes	Embrun	Embrun	clerc de notair	employé	68	31,2	12	oui
23	Alpes Hautes	Embrun	Embrun	cultivateur	agrindép	64	31,2	0	oui
24	Alpes Hautes	Embrun	Embrun	cultivateur	agrindép	70	31,2	1	oui
25	Alpes Hautes	Embrun	Embrun	cultivateur	agrindép	39	31,2	12	oui
26	Alpes Hautes	Embrun	Embrun	cultivateur	agrindép	49	31,2	12	oui
27	Alpes Hautes	Embrun	Embrun	cultivateur	agrindép	50	31,2	12	oui