data.ameli.fr, France's national health insurance open data.

France's statutory health insurance, the Assurance Maladie, reimburses care for almost the entire population. data.ameli.fr is the open data portal of its national fund, the Cnam: tables of totals built from those reimbursements, free and without registration. **The portal, the field names and the category labels are in French only.**

1 · COLLECTION
Reimbursements
Every reimbursed item of care (consultation, drug, hospital stay, sick leave) creates a record. 67.6 M people covered in 2024.

→

2 · INDIVIDUAL DATABASE
The SNDS, closed
The Système national des données de santé links each person's care over time. Access on authorisation and protocol only; care pathways stay there, out of the portal's reach.

→

3 · CASE DEFINITION
The Cnam algorithm
The Cartographie des pathologies assigns diseases from ALD status, hospital diagnoses and tracer drugs. The algorithm is versioned: two releases do not join up.

→

4 · OPEN
data.ameli.fr
Aggregate counts by year, area, sex and age. 51 datasets, ≈ 13.0 M rows, API without authentication, ODbL licence.

Consequence of step 3: the central indicator is a **treated prevalence**, the people their reimbursements make visible. Someone ill but undiagnosed or not followed appears nowhere. And when a figure rises, three causes produce the same effect and cannot be told apart: more new cases, more screening, or better access to care.

FOUR DOMAINS, JOINABLE BY YEAR AND, FOR MOST, BY AREA (NO INDIVIDUAL KEY)

Diseases
3 DATASETS · 2015-2024 · 5.9 M ROWS

`year × disease (60 categories, 3 levels) × sex × age × area → ntop, npop, prev · + spending, comorbidities`

Prevalence, cost per patient and per spending item, comorbidities. Cancers separate active treatment and follow-up.

EXAMPLE
Diabetes prevalence more than doubles across regions: 4.9% in Brittany, 11.9% in Guadeloupe (2024, all ages). Age being a dimension, the comparison can be redone at equal age.

Private-practice professionals
8 DATASETS · 2010-2025 · 5.9 M ROWS

`year × profession (38) × area (× age × sex) → headcount, density, fees, prescriptions`

Medical demography, extra billing (the effect of the Optam contract can be isolated), prescriptions by item.

EXAMPLE
The density of private-practice GPs varies by a factor of 2.6 in mainland France: 48 per 100,000 inhabitants in Seine-Saint-Denis, 127 in Hautes-Alpes (département level, 2025).

Access to care
16 DATASETS · 2012-2025 · 77 K ROWS

`year × area (× profession) → value, annual change · infra-annual reading at 30 June 2026`

8 public policy indicators: regular doctor, first-time practice openings, under-served areas, care access service.

EXAMPLE
Coverage of the care access service rises from 29% of the population in 2021 to 98.5% in 2025: a complete roll-out curve, usable as a treatment variable.

Occupational risks
16 DATASETS · 2005-2024 · 1.1 M ROWS

`year × business sector (NAF) or region × risk → claims, employees, frequency index, recognised diseases`

Work and commuting accidents, recognised occupational diseases, sick leave, permanent disability and deaths, by business sector and by region.

EXAMPLE
Musculoskeletal disorders account for 88% of the 50,598 occupational diseases recognised in 2024.

GETTING THE DATA

Web interface · exploration, filters, map, export. Shows the exact labels of each category, in French.

File exports · 10 formats, no row limit. Main dataset: ≈ 1 GB as CSV, 50 MB as Parquet. Same filters as the API.

Explore API v2.1 · ODSQL, server-side aggregation, no authentication. 100 rows per call, pagination window of 10,000 (20,000 when aggregated).

data.gouv.fr MCP server · no key, one filter criterion per query; covers catalogue exploration.

A TYPICAL QUERY

```
GET /api/explore/v2.1/catalog/datasets/effectifs/records
 ?where=year(annee)=2024 AND region=99 AND dept="999"
   AND sexe=9 AND cla_age_5="tsage"
   AND niveau_prioritaire=1
 &select=patho_niv1, sum(ntop) as n
 &group_by=patho_niv1 &order_by=n desc
```

Disease families ranked for France as a whole, 2024, computed server-side. Field names are French: `annee` year, `sexe` sex, `cla_age_5` five-year age band, `patho_niv1` disease family.

BEFORE YOU START

1 — Totals inside the table
Aggregate rows (`region=99` + `dept="999"` , `sexe=9` , `tsage`) sit alongside the detail in the same tables.

2 — Silent failures
An unrecognised parameter is ignored: response 200, unfiltered table. A filter that did not apply shows in the response's `total_count` .

3 — Diseases on 3 levels
The table holds all three levels of the tree (family, disease, detail): `niveau_prioritaire` selects the one you read. One person can fall into several categories.

4 — Statistical confidentiality
Cells under 11 patients are blanked, 26% of the main dataset (threshold of 5 for professionals). Deterministic masking, correlated with the value.

5 — Mixed typing
`prev` is a precomputed ratio. `annee` is typed as date or text depending on the dataset; text and `_integer` duplicates, sometimes on different scales.

Not covered: incidence (Francim cancer registries), mortality (CépiDc), individual care pathways (SNDS, on authorisation), areas below the département, cancer stage and treatment.