

VI Digital Government Ministerial Meeting of the Americas

OPEN DATA THE FRENCH CASE

Antonin Garrone

Mario Restuccia

Wednesday 18 November 2020

Who we are



Antonin Garrone

Data Editor



Mario Restuccia

Open Data Officer

MISSION

What is Etalab

We coordinate the design and implementation of the State's data strategy.

What is Etalab

- Data **opening** and data circulation
- Data **exploitation**: data science, artificial intelligence
- Innovation and openness of **public action**



Our programme today

Why?

The **importance**
of public Open
Data

Case studies

How?

Open Data Policy
Open Data
Implementation
Open Data
Quality

In practice

The impact of
COVID-19

**Why Open
Data?**

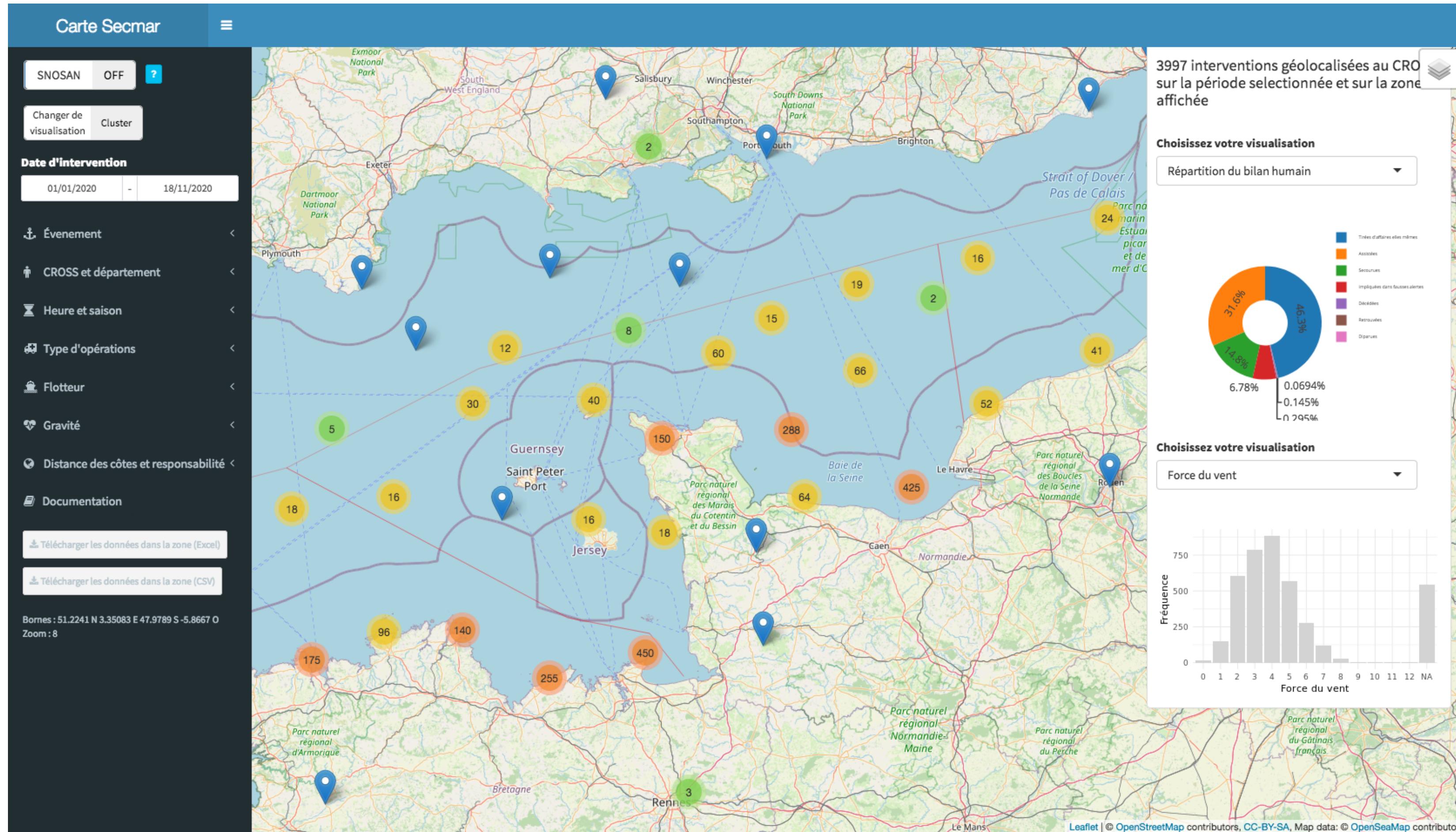
WHY OPEN DATA?

Foster public service transformation

- **Gain of efficiency:** administrations don't have to respond to individual requests
- **Improve the quality of public service:** rely on reference data
- **Data-driven public policies**

WHY OPEN DATA?

Reducing accidents at sea



carte.snosan.fr

WHY OPEN DATA?

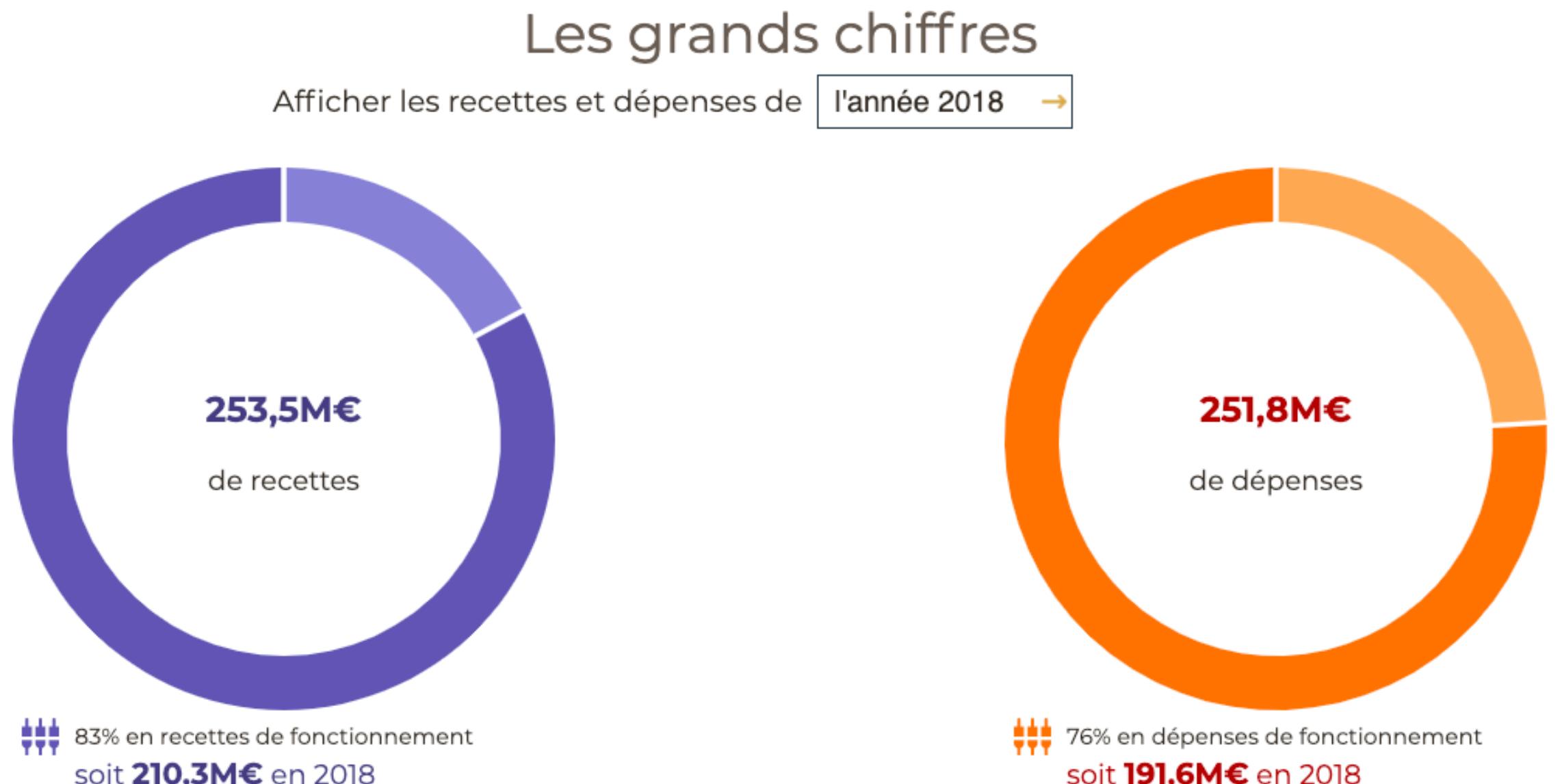
Build trust

- Facilitating the **understanding of public spending or public procurement**
- Effectively **informing citizens** in times of crisis

WHY OPEN DATA?

Transparency of public expenditures

Explorer les comptes de la ville



83% en recettes de fonctionnement
soit **210,3M€** en 2018

Les **recettes de fonctionnement** regroupent l'ensemble des ressources nécessaires au fonctionnement des services de la ville. Ce sont **principalement des recettes fiscales**.

► Afficher des exemples de recettes.

76% en dépenses de fonctionnement
soit **191,6M€** en 2018

Les **dépenses de fonctionnement** comprennent toutes les dépenses nécessaires au fonctionnement des services de la ville, c'est-à-dire les **dépenses qui reviennent régulièrement chaque année**.

► C'est-à-dire ?

WHY OPEN DATA?

To enable new services

- Data as an **infrastructure**
- Making it easier for citizens to **get around**
- Fluidify the **real estate market**

WHY OPEN DATA?

Fluidify the real estate market



app.dvf.etalab.gouv.fr

French Open Data Policy

The Digital Republic Act

- Principle of "**Open Data by default**"
- **Obligation** for local governments
- Private actors in charge of a public service mission
- Specific **sectoral legislation** (waste, public procurement, etc.)



Public Service Data

- Notion of "**reference data**"
- "**Guaranteed level of service**" (**data availability, freshness and quality**)
- Nine datasets at the moment

GENERAL FRAMEWORK

Public Service Data

Les données de référence

À ce jour, neuf jeux de données, qui couvrent un large champ thématique ont été identifiés comme des données de référence.



Base Adresse Nationale (BAN)

La Base Adresse Nationale est l'une des neuf bases de données du service public des données de référence. Elle est la seule

1 33 41 France



Base Sirene des entreprises et de leurs établissements (SIREN, SIRET)

Pour vous abonner à notre lettre d'information Sirene open data actualités, suivez ce lien :

15 7 30 France



Code Officiel Géographique (COG)

Chaque année, l'Insee met à disposition sur son site (insee.fr) le code officiel géographique qui rassemble les codes et

35 13 24 France Commune française



Plan Cadastral Informatisé (PCI)

Le plan cadastral est un assemblage d'environ 600 000 feuilles ou planches représentant chacune une section ou une

3 10 57 France



Registre parcellaire graphique (RPG) : contours des parcelles et îlots culturels et leur groupe de cultures majoritaire

Le registre parcellaire graphique est une base de données géographiques servant de référence à l'instruction des aides de la

21 2 13 2013–2017 France



Référentiel de l'organisation administrative de l'Etat

Le décret du 14 mars 2017 a institué le Service Public de la Donnée. Celui-ci met à la disposition du public 9 jeux de données

5 4 16 2016–2026 France



Référentiel à grande échelle (RGE)

L'État a confié à l'IGN le développement du référentiel à grande échelle (RGE). Pour ce faire, il fait appel à ses moyens propres

12 2 8 France



Répertoire National des Associations (RNA)

** ! Des emails frauduleux usurpant l'identité de data.gouv.fr circulent à destination des associations en demandant

27 7 29 1901–2017 France Point



Répertoire Opérationnel des Métiers et des Emplois (ROME)

Dans un contexte marqué par de fortes mutations de l'environnement économique et social, le ROME (Répertoire Opérationnel

5 2 9



Open Data implementation

data.gouv.fr

Open platform for French public data



Search

Agriculture et Alimentation

Culture, Communications

Comptes, Économie et Emploi

Éducation, Recherche, Formation

International, Europe

Environnement, Énergie, Logement

Santé et Social

Société, Droit, Institutions

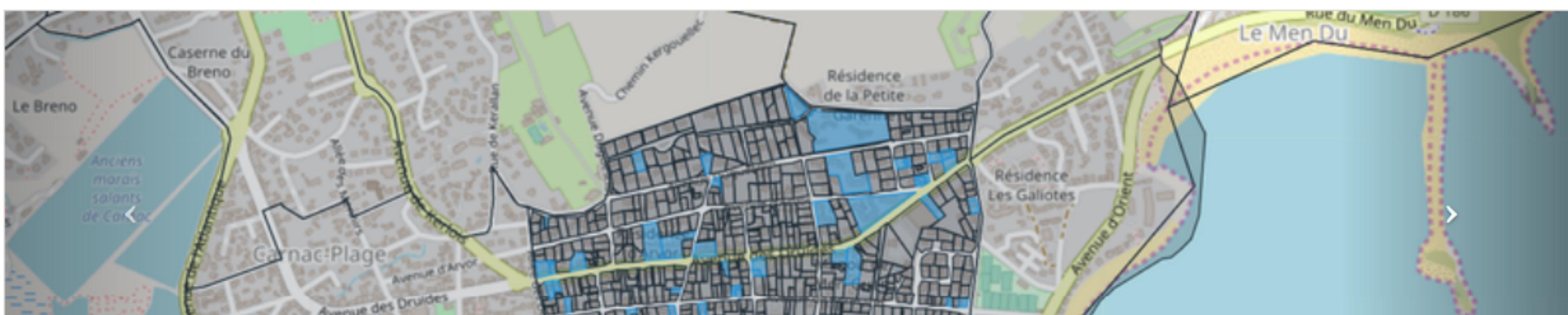
Territoires, Transports, Tourisme

Share, improve and reuse public data



CONTRIBUTE!

BEST REUSE CASES



ETALAB

EXPLORATEUR DE DONNÉES DE VALEUR FONCIÈRE (DVF)

DATA.GOUV.FR

35 000+

Datasets

2 400+

Reuses

2 500+

Publishers

LITERACY

Assisting and empowering producers

- Foster an **open data culture** among the administration
- Providing **technical support**
- Helping producers with **legal matters**

Guides & docs

 guides.etalab.gouv.fr

- Aimed at chief data officers, lawyers, project managers, or just users
- Open source, of course
 [etalab/guides.etalab.gouv.fr](https://github.com/etalab/guides.etalab.gouv.fr)

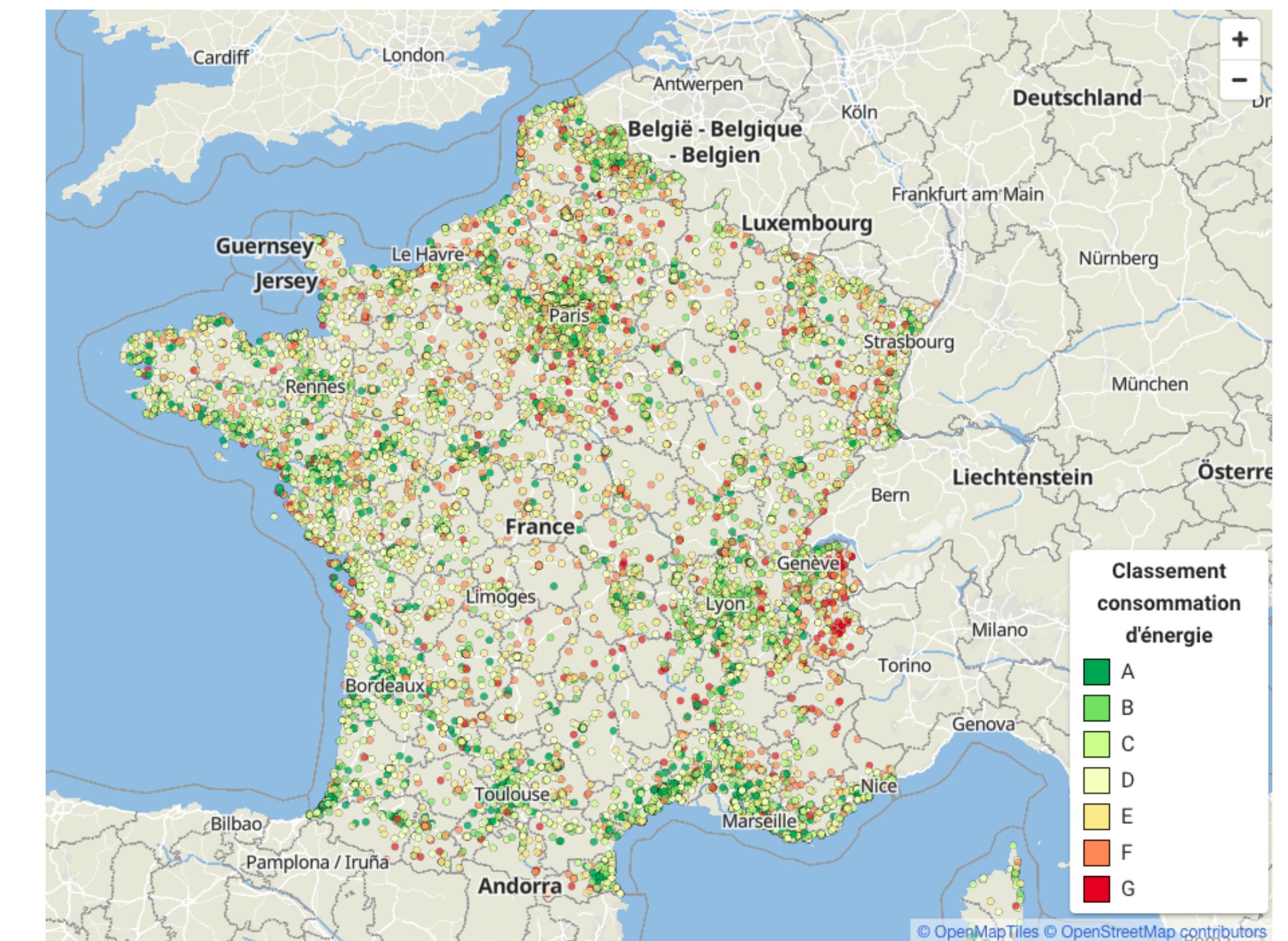
 doc.data.gouv.fr

- Documentation on how to publish data on the portal
- Also open source
 [etalab/doc.data.gouv.fr](https://github.com/etalab/doc.data.gouv.fr)

DISCOVERABILITY

Community animation

- Foster **exchanges** and promote opening and reuses
- DataFin hackathon on local government financial data held the French Senate
- RenovAction hackathon dedicated to energy-efficient renovation of buildings



DISCOVERABILITY

Curation & editorial



Curation

- Editorialisation and **curation** work on the data.gouv.fr portal
- Inventory pages



Content

- Monthly article to **showcase** the most important publications
- Specific article on **high potential datasets**

**open Data
quality**

QUALITY

Metadata monitoring

When creating a dataset on data.gouv.fr, producers are **encouraged** to provide metadata

Quality



The aim of that box is to help you improve the quality of the (meta)data associated to your dataset.

It gives you an overview of what will be useful for contributors to find and reuse your data.

! Description

! Tags

✓ Open formats

? Discussions

? Up-to-date

✓ Availability

Data schemas

- A schema is a **blueprint** of how the data should be structured
- They allow an increase in the **quality** of the data, especially when several data producers create the same kind of data



```
{  
  "fields": [  
    {  
      "name": "Name",  
      "type": "string",  
      "description": "User's name"  
    },  
    {  
      "name": "Email",  
      "type": "string",  
      "format": "email",  
      "description": "User's email"  
    },  
    {  
      "name": "Age",  
      "type": "integer",  
      "description": "User's age",  
      "constraints": {  
        "minimum": 18  
      }  
    }  
  ]  
}
```

Name	Email	Age
Jill	jill@foo.com	25
Jack	jack@bar.com	33

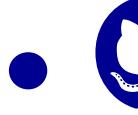
QUALITY

Utilities for data schemas

 [validata.etalab.studio](#)

- Check if data conforms to a schema
- A shared project among different actors
↳ git.opendatafrance.net/validata

 [csv-gg.etalab.studio](#)

- Generate a csv that conforms to a schema
-  [etalab/csv-gg](#)

Open Data in practice : the case of COVID-19

COVID-19

The reason why

- Provide factual **information** to citizens and policy makers
- Essential for researchers, journalists, companies, etc.

Different types of data

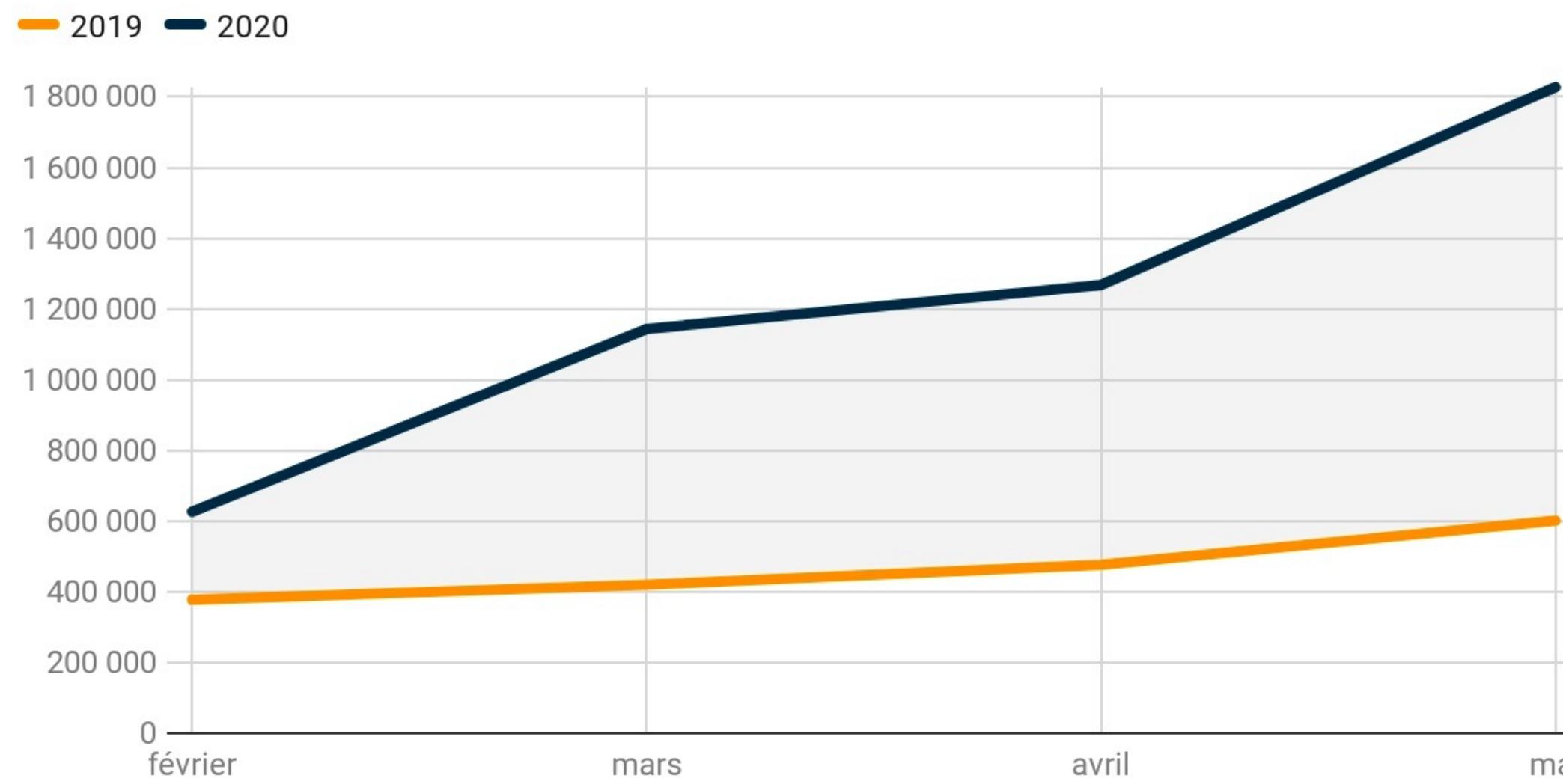
- **Health** data (hospital and ICU capacity, testing, excess mortality, etc)
- **Economic** impact
- But also education, research, etc.

COVID-19

What we did

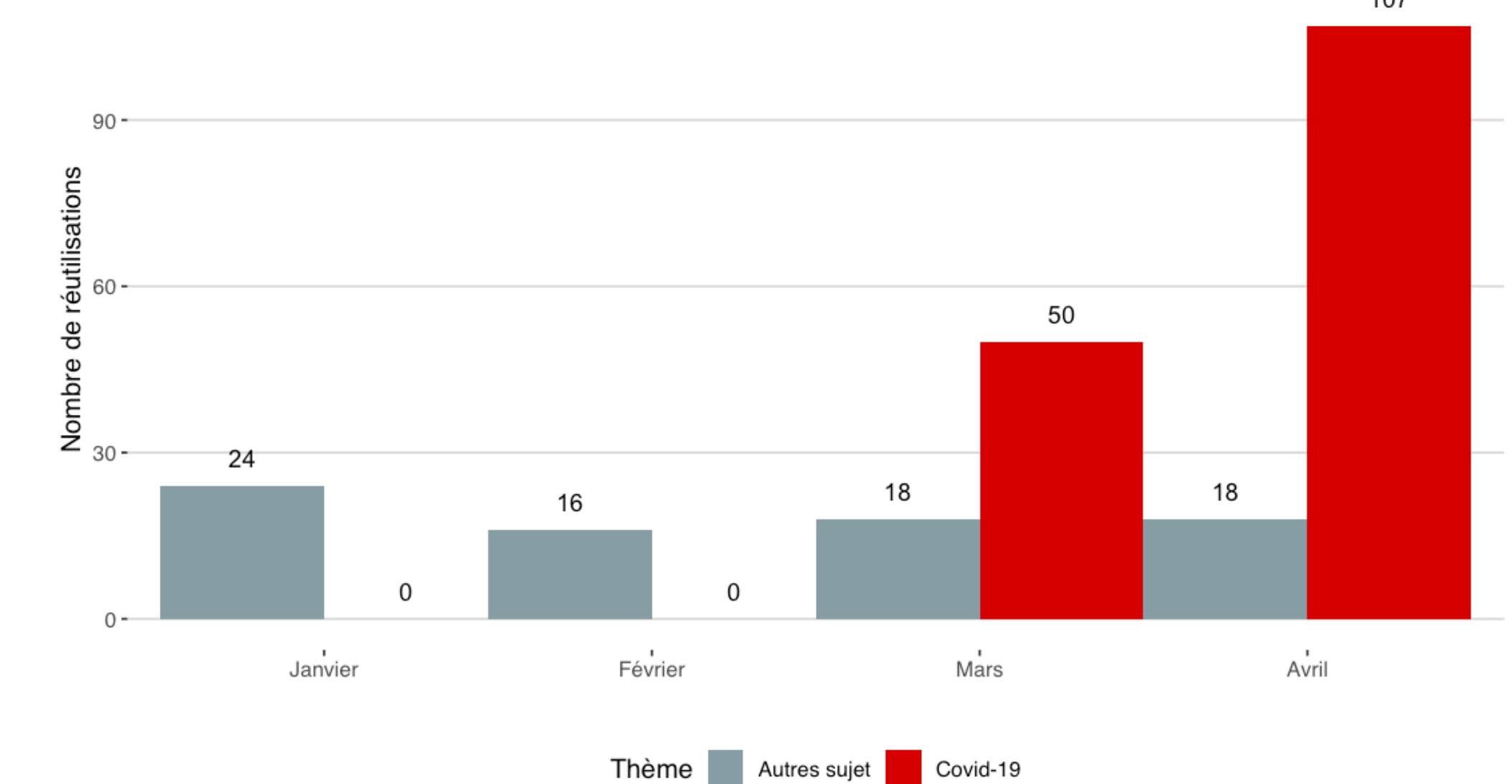
Absolute record in traffic

Évolution du trafic sur data.gouv.fr entre février et mai 2020 par rapport à février-mai 2019



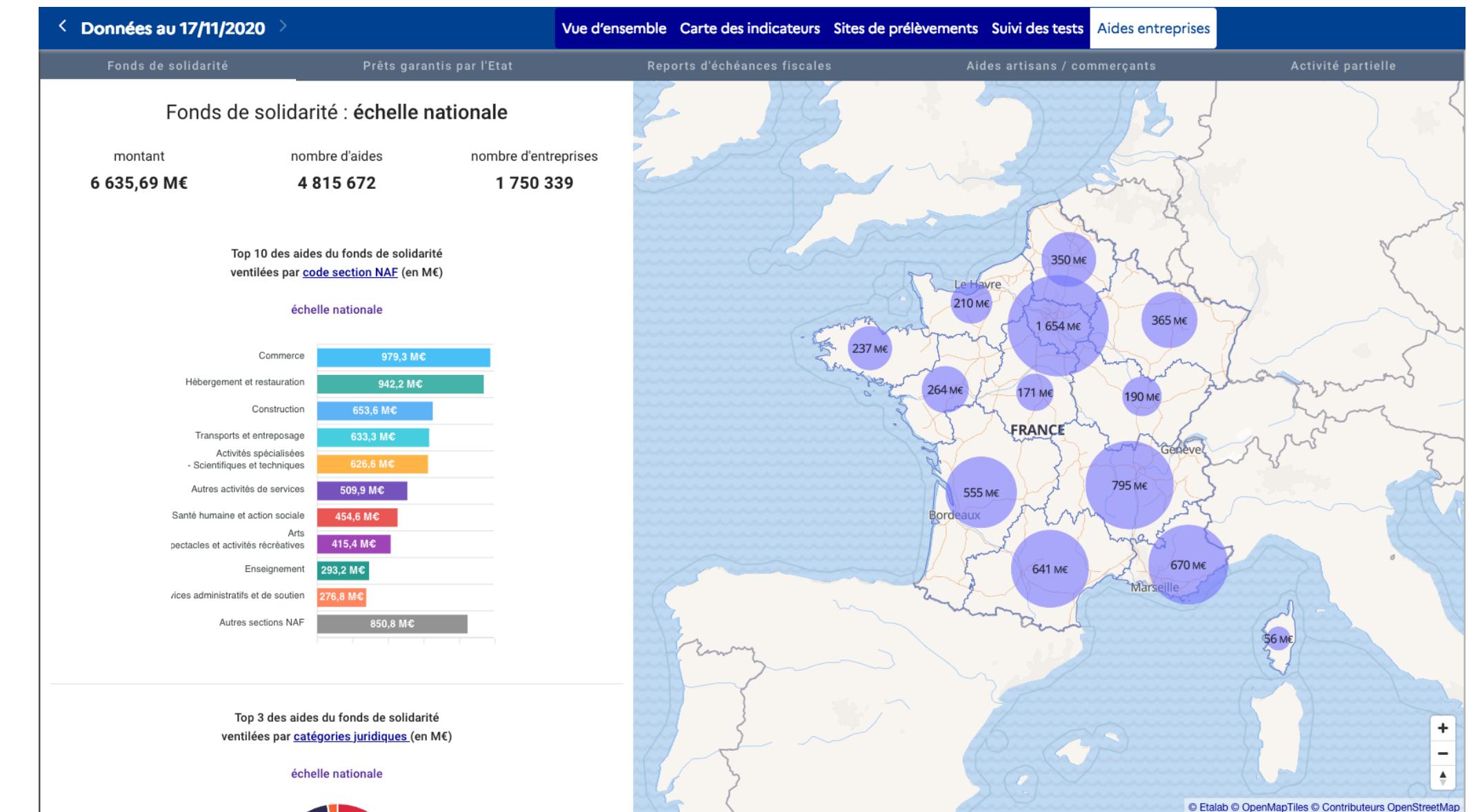
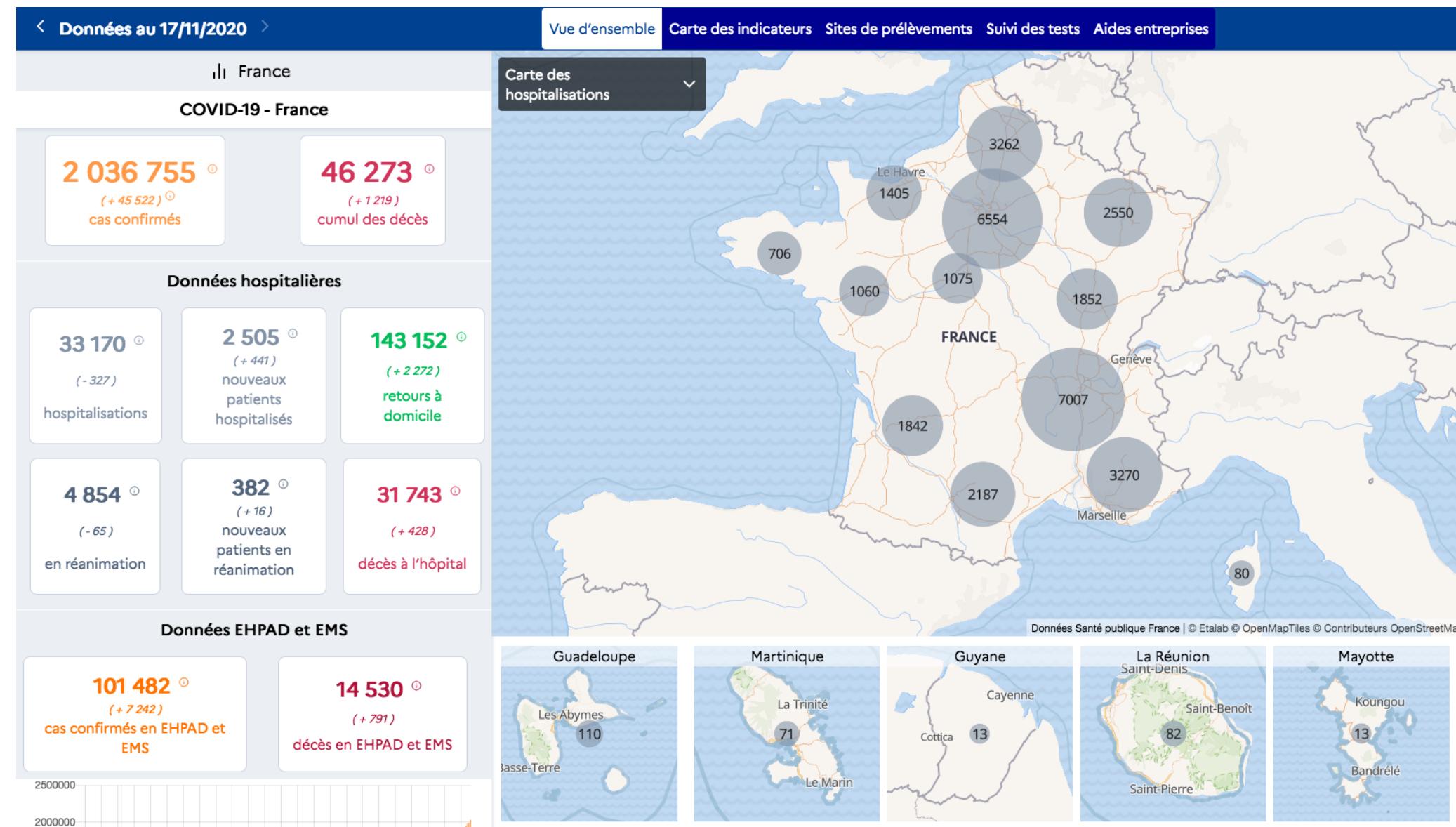
And numbers of reuses

Impact des données relatives au coronavirus sur le nombre de réutilisations référencées sur data.gouv.fr



COVID-19

Make data discoverable and understandable



Monitoring the epidemic

Monitoring support to companies



gouvernement.fr/info-coronavirus/carte-et-donnees

COVID-19

Several challenges

Organisational challenge

Coordination among actors

Technical challenge

Different data sources

Evolving situation

Need to adapt the indicators

COVID-19

What we learned

Be prepared

implement interoperable systems of data collection, aggregation and publication

Be broad

data usually considered to be of low priority can become **critical** in a crisis

Be organised

map all the **data** produced and associate operational contacts for each dataset

**Ambitions for
the future**

A renewed (political) interest

Lessons from crisis management

- COVID-19 raised the salience of data in the political debate
- Policy decisions must be backed by data

FUTURE

A renewed (political) interest

Through the Recovery Plan, the Government will finance data-driven projects within public administrations

Data-driven decision-making

Data circulation through APIs

Capitalising on existing solutions (e.g. already available data)

MERCI !

 antonin.garrone@data.gouv.fr

 mario.restuccia@data.gouv.fr

Wednesday 18 November 2020

RESOURCES

 _DINUM

 etalab

 datagouvfr

 etalab

 etalab

 numerique.gouv.fr

 etalab.gouv.fr

 data.gouv.fr