

INTRODUCTION ET ENJEUX DE L'OPEN DATA À SANTÉ PUBLIQUE FRANCE

Yann Le Strat Les rencontres de Santé publique France 25 mai 2021



Cette intervention est faite en tant que personnel de Santé publique France, organisateur de la manifestation.

Je n'ai pas de conflit d'intérêt en lien avec le sujet traité.

DÉBUT DE L'OPEN DATA À SANTÉ PUBLIQUE FRANCE



- L'open data a été effectif le 1^{er} février 2019 à Santé publique France.
- Conférence de presse et communiqué de presse

Géodes, une avancée dans la restitution des données de surveillance

L'une des missions de Santé publique France est de produire des indicateurs utiles à la décision et à l'action en santé publique, en s'adaptant aux évolutions et aux enjeux sociétaux, numériques et scientifiques. Jusqu'à présent, la restitution des indicateurs se faisaient par l'intermédiaire de bulletins ou de publication scientifiques. L'observatoire Géodes répond à une volonté forte de transparence et de modernité en permettant un accès libre simplifié à de nombreux indicateurs de l'agence.

Les utilisateurs pourront grâce à cette plateforme, accessible en ligne :

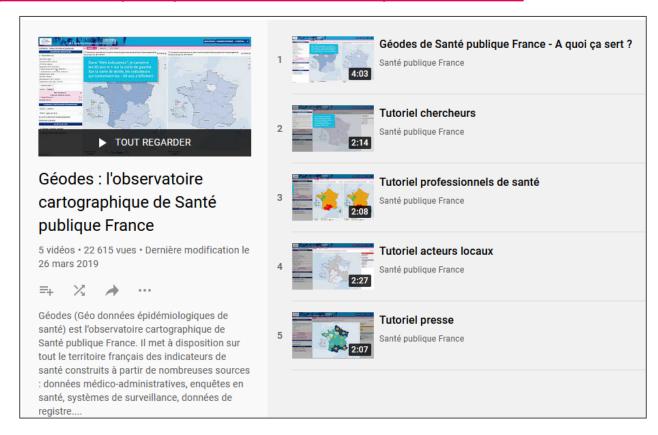
- Accéder à plus de 300 indicateurs concernant 40 thématiques
- Faire des comparaisons entre des indicateurs régionaux et nationaux
- Avoir le détail technique de chaque indicateur : la source, la période de mise à jour ...
- Télécharger des cartes et des données et les partager sur les réseaux sociaux
- Etre orienté pour chaque indicateur vers la page thématique du site internet de Santé publique France
- Importer des données dans géodes et les comparer à celles de Santé publique France.

OUVERTURE DE GÉODES ET DES TUTORIELS VIDÉO



Ouverture du site Géodes :

https://geodes.santepubliquefrance.fr/#view=map2&c=indicator



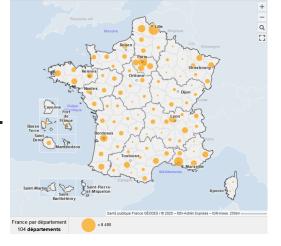
https://www.youtube.com/playlist?list=PLI00syIAMv7TLpIsDMY06pEUgQwYt5cxj

OBJECTIFS DE GÉODES



 Rendre accessible nos indicateurs à des personnes travaillant plutôt dans le domaine de la santé publique, même si tout un chacun peut y accéder.

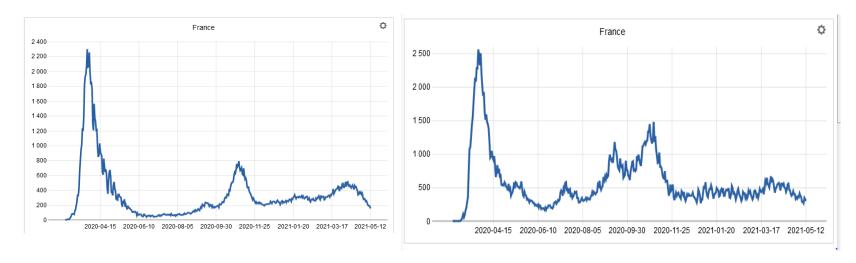
- C'est un « observatoire » qui propose un outil cartographique, un export des données, des séries temporelles, un formulaire de contact, etc.
- Ce n'est pas un tableau de bord!
- 300 indicateurs épidémiologiques au lancement
- Outil qui centralise tous les indicateurs de l'agence (infectieux, maladies chroniques, santé environnementale, au travail, déterminants)
- 90 000 visites en 2019 et puis est arrivée l'épidémie de Covid-19



EN FÉVRIER/MARS 2020



 A partir du 24 février, mise en ligne sur Géodes du taux de passage aux urgences pour Covid-19, grâce au système de surveillance Oscour® et du taux d'actes codés Covid-19 par SOS-Médecins.



 Le 17 mars 2020, les premières données concernant les hospitalisations liées à la Covid-19 arrivent et sont traitées par l'agence.

17/18 MARS 2020



Réunion avec le Service d'Information du Gouvernement (<u>SIG</u>) et <u>Etalab</u>, département de la Direction Interministérielle du Numérique (DINUM), qui coordonne la politique d'ouverture et de partage des données publiques, développe et anime la plateforme d'open data <u>data.gouv.fr</u>

- Décision immédiate que :
 - Santé publique France transmette des fichiers d'indicateurs à Etalab -> mise en place d'un SFTP
 - data.gouv.fr sera le point central de stockage des fichiers
 - Création d'une page de Santé publique France le 18 mars https://www.data.gouv.fr/fr/organizations/sante-publique-france/
 - A partir de cette date, tous les indicateurs liés
 à la Covid-19 produits par Santé publique France
 sont transmis quotidiennement (7j/7) par ce canal.



LES TRANSMISSIONS D'INDICATEURS





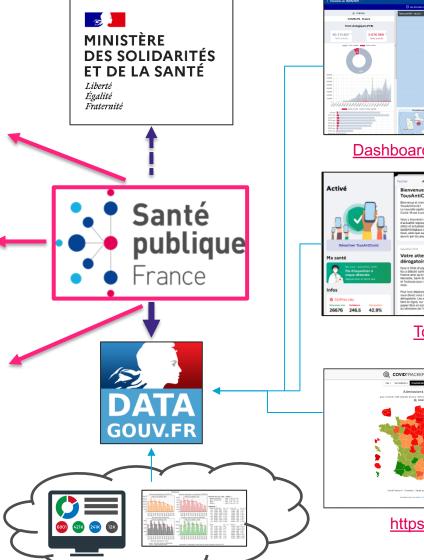
<u>Géodes</u>



Site de Santé publique France



<u>European Centre for Disease Prevention</u> and Control

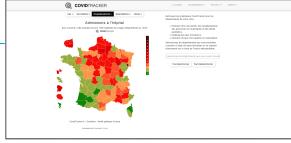




Dashboard site du gouvernement



TousAntiCovid



https://covidtracker.fr/

GÉODES EST TRÈS UTILISÉ PAR LA PRESSE LOCALE



a Actu.fr

Oise : le taux d'incidence du Covid-19 en dessous de 270 ...

Depuis ce début de semaine, le site Géodes affiche dans l'Oise un taux d'incidence en dessous de 270 pour 100 000 habitants. Baisse du taux ... Il y a 1 jour



vm Var-Matin

Covid-19: la carte de l'épidémie dans le Var, commune par ...

On fait le point sur l'activité du virus dans le Var, commune par commune, après trois semaines de confinement. Capture d'écran / Géodes.



C E-santé

Thrombose veineuse, embolie pulmonaire : les 10 ...

... femmes, Géodes https://geodes.santepubliquefrance.fr/#c=indicator&f=0& i=cardio_mvte.mvte_txstd&s=2019&selcodgeo=08&t=a01&view= ...
Il y a 2 jours



a Actu.fr

Covid-19 en Bretagne : voici les communes où circule le plus ...

Mais aussi le moins ? On fait le point avec l'outil Geodes sur le taux d'incidence entre le 6 avril et le 12 avril 2021.

Il y a 1 mois



C'Internaute

Variant du Covid en France : quelles sont les derniers chiffres ? La Carte

Depuis le lundi 8 mars, Santé publique France actualise sur son site Géodes plusieurs cartes affichant la progression des variants anglais, ...

Il y a 1 jour



BP Le Bien Public

Santé. Covid-19 : où en est la campagne de vaccination en ...

... accélération de la campagne de vaccination dans le département ? Oui, a priori, si l'on se fie aux chiffres de Geodes/Santé publique France.

II y a 2 jours



La Tribune

Coronavirus: Légère hausse du nombre de patients en ...

Selon les chiffres disponibles sur le site Géodes du ministère de la Santé, 145 nouveaux décès ont été recensés a l'hôpital en raison du ...

II y a 2 semaines



nm Nice-Matin

La carte de l'épidémie de Covid-19 dans les Alpes-Maritimes ...

Pour la semaine du 14 au 20 avril février (les données les plus récentes publiées sur la plateforme gouvernementale Géodes), le taux ...

II y a 3 semaines



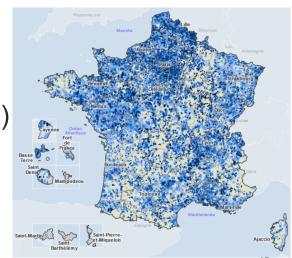
FRÉQUENCE ÉLEVÉE



En 2020 : 31 millions de visites, 107 millions de pages vues



- Très forte fréquentation lors de l'ouverture des indicateurs à des niveaux infra-départementaux (EPCI, Communes, IRIS)
- En 2021 : 12 millions de visites, 40 millions de pages vues.



QUELS ENSEIGNEMENTS?



- Très grande visibilité avec Géodes, data.gouv, dashboard gouvernement, divers dashboards (data scientists) + sites internet (notamment de la presse)
- La France est leader dans la réutilisation des données d'open data de la Covid-19 (enquête auprès de 52 pays) avec 129 applications et a été dans les plus rapides à ouvrir ses données après le 1er cas.
- Les Français avaient accès à toute l'information, en temps réel, en transparence, dès le début de l'épidémie

#	Country	Release date & frequency of updates [29]	machine- readability [29]	1st case of Covid-19 [7]	Open data available	I st mentioning (DD/MM/YYYY)	Release date (DD/MM/2020)	# of data sets	FC/ODA (-1,0,1)	EDP available #	frequently? (yes/no)	machine- readable? (yes/no)
1	France	3,00	2,97	24.02.	1	09.09.2014	29.02.	106	1	362	yes	yes
2	Spain	2,91	2,09	31.01.	1	26.06.2018	n/a, March	97	0	108	yes	yes
3	Ireland	2,55	2,36	29.02.	1	31.03.2020	31.03.	11	-1	12	yes	mainly
4	Cyprus	2,84	2,69	09.03.	1	18.12.2017	07.04.	19	0	17	yes	yes
5	Finland	2,85	2,24	29.01.	0							
6	Slovenia	3,00	2,94	04.03.	0							
7	Austria	2,97	2,84	25.02.	1	25.02.2020	25.02.	16	1	1	no	yes
8	Romania	2,13	2,03	26.02.	0							
9	Luxembourg	2,91	2,00	29.02.	1	02.04.2020	23.04.	29	-1	30	yes	yes
10	Netherlands	2,13	2,00	27.02.	1	n/a	n/a	12	n/a	16	yes	not always
11	Latvia	2,90	2,10	02.03.	1	27.03.2020	27.03.	2	0	2	yes	yes
12	Poland	2,94	2,00	04.03.	1	20.05.2020	-/- 01 02	1 29	0	-	no	yes
13	Italy	2,15	2,18	30.01.	1	01.03.2020	n/a, 01.03.			21	yes	yes
14	Greece.	2,88	2,00	27.01. 26.02.	1	25.03.2020	25.03.	11	-1	35	yes	yes
15		2,28	2.03	25.02.	0	28.04.2020	28 04	١.				
16	Croatia Belgium	1,94	2,03	04.02.	1	n/a	28.04. n/a	1 17	n/a	1 16	yes	yes
17 18	Estonia	3.00	2.06	27.02.	1	14.03.2020	14.03.	3	11/a	10	yes	yes
18	Denmark	1.19	2.06	27.02.	1	16.03.2020	14.03.	1	U		yes no	yes
20	Norway	2,19	2,00	26.02.	0	10.05.2020	-	1		-	110	no
21	Bulgaria	2,19	2,10	08.03.	0							
22	UK	2,20	2.00	31.01.	1	07.04.2020	05.06.	17	-1	29	not always	mainly
23	Malta	1,97	1.66	07.03.	0	07.04.2020	03.00.	17		27	not arways	шашту
24	Switzerland	2,63	2.09	25.02.	1	05.11.2019	05.03.	7	1	5	ves	mainly
25	Portugal	2,63	2.00	02.03.	0	05.11.2017	03.03.	· ·	•	1	yes	mounty
26	Sweden	2,16	2.00	31.01.	1	27.03.2020	27.03.	5	-1	4	yes	yes
27	Lithuania	1.84	2.16	28.02.	1	10.04.2020	10.04	1	-1		no	yes
28	Czech Republic	1,63	2,00	01.03.	1	n/a	n/a	16	n/a	18	yes	yes
29	Slovakia	2,79	2,03	06.03.	0							
30	Hungary	1,81	2,06	04.03.	0							
31	Iceland	2,53	2,83	28.02.	0							
32	Russia	2,19	2,90	31.01.	0							
33	Taiwan	2,78	2,31	21.01.	1	12.04.2020	12.04.	6	-1		110	yes
34	Canada	2,97	2,16	25.01.	1	06.02.2020	15.04.	62	-1		yes	not alway
35	Colombia	2,67	1,97	06.03.	1	27.03.2020	27.03.	6	0		yes	yes
36	New Zealand	2,06	2,03	28.02.	1	02.03.2020	-	13	-		no	yes
37	India	2,13	2,91	30.01.	1	07.05.2020	07.05.	156	-1		not always	yes

Nikiforova A (2020) <u>Timeliness of open data in open government data portals through pandemic-related data: a long data way from the publisher to the user</u>. Fourth international conference on multimedia computing, networking and applications (MCNA).

EN TANT QUE PRODUCTEUR (1/2)



- Assurer la non identification des personnes à partir des indicateurs produits!
- Accepté d'être challengé : exposition forte, nombreuses utilisations, nombreuses questions, commentaires (élogieux ou non) en lien ou non avec ce qui est produit.
- En bénéficier :
 - amélioration des méthodes,
 - devenir très pointu sur la formulation, bien faire comprendre les termes utilisés et leurs différences,
 - corriger des non mises à jour ou des petits bugs

L'ouverture des données et des méthodes protège le producteur

EN TANT QUE PRODUCTEUR (2/2)



Prévoir un accompagnement à l'ouverture des données

- En termes de communication :
 - Expliquer ce qui va être mis en ligne (tutelles, data.gouv, TAC, communiqués de presse, réunions avec des utilisateurs potentiels de nouveaux fichiers, datajournalistes, etc.)
 - Expliquer la méthodologie (note méthodologique, réunions de présentation)
 - Fournir des méta-données, notices citant la source des données, points clés.
- En termes de réponse aux sollicitations :
 - Demande d'explications (journalistes+++, équipes de recherches, grand public)
 - Sollicitations trop nombreuses pour la Covid-19 pour pouvoir y répondre efficacement.
 - Ouvrir ses données nécessite un « service après vente » à anticiper !

LES VERTUS DE L'OUVERTURE DES DONNÉES (RAPPORT DE LA MISSION BOTHOREL)



La donnée est :

- Un vecteur de connaissance (science ouverte)
- Un levier d'innovation
- Améliore le service public en interconnectant la puissance politique et l'usager
- Est un moyen puissant de restaurer la confiance dans l'action publique



Politique, la donnée est un moyen puissant de restaurer la confiance dans l'action publique. Le défi est immense : la consultation publique conduite par la mission a révélé une défiance majeure dans la donnée utilisée par l'État. À cet égard, la crise de la Covid19 a provoqué un éveil de la nation à la donnée. Les difficultés pour établir des statistiques en temps réel de l'épidémie, les conflits d'interprétation des courbes, la fragilité des hypothèses des modèles épidémiologiques, auront eu une vertu pédagogique : celle de révéler que toute donnée est une construction et qu'elle doit être expliquée. Ouvrir la donnée, c'est enrichir le débat public. Cette transparence est le meilleur remède à la défiance et au complotisme.

https://www.gouvernement.fr/remise-du-rapport-sur-la-politique-publique-de-la-donnee-des-algorithmes-et-des-codes-sources

DISCUSSION



- L'ouverture des données est nécessaire et justifiée pour une agence publique et scientifique
- Agence publique : les données sont collectées sur des fonds publics et il est légitime de les exploiter au mieux et de les restituer
- Agence scientifique : transparence, ouverture, communication des résultats
- Au service de la population :
 - Chaque français doit disposer de toute l'information, au même moment.
 - Le maillage géographique très fin pour les indicateurs de la Covid-19 avait aussi pour but d'accroitre la prise de conscience des enjeux liés à l'épidémie
 - Rendre plus concret le niveau épidémique à l'échelle d'une ville ou d'un quartier
 - Faciliter la compréhension et l'acceptation de la population aux mesures priss pour lutter contre l'épidémie.

REMERCIEMENTS



- Merci à toutes les personnes ayant contribué à l'ouverture des données, notamment celles du Covid-19 : Adel, Agnès, Aurélie D, Aurélie F, Delphine, Etienne, Jebraiel, Jessica, Marc, Perrine
- Merci à Etalab pour l'excellente coopération : remerciements spéciaux à Mathilde Hoang et à son équipe.
- Merci à Emc3 (Eric Mauvière et son équipe), à Ciril GROUP (Anthony Legros et ses collègues)
- Et merci à toutes les personnes de ma direction ayant construit tous les indicateurs en amont, chaque jour depuis plus d'un an !