ESPON

EUROPEAN UNION

Co-financed by the European Regional Development Fund

Inspire Policy Making with Territorial Evidence

ESPON STUDY //

Geography of COVID-19 outbreak and first policy answers in European regions and cities

Annex : data collection process // December 2020

This ESPON STUDY is conducted within the framework of the ESPON 2020 Cooperation Programme, partly financed by the European Regional Development Fund.

The ESPON EGTC is the Single Beneficiary of the ESPON 2020 Cooperation Programme. The Single Operation within the programme is implemented by the ESPON EGTC and co-financed by the European Regional Development Fund, the EU Member States and the Partner States, Iceland, Liechtenstein, Norway and Switzerland.

This delivery does not necessarily reflect the opinions of the members of the ESPON 2020 Monitoring Committee.

Coordination:
Sebastien BOURDIN, EM Normandie Business School
Andreea CHINA, ESPON EGTC

Authors
Sebastien BOURDIN, EM Normandie Business School
Gabriel NOIRET, EM Normandie Business School

Acknowledgements
EUROCITIES, Council of European Municipalities and Regions (CEMR), European Confederation of Local Intermediate Authorities (CEPLI), ESPON contact points across Europe.

Information on ESPON and its projects can be found at www.espon.eu.
The website provides the possibility to download and examine the most recent documents produced by finalised and ongoing ESPON projects.

ISBN: Click or tap here to enter text.

© ESPON, 2020
Published in Click or tap here to enter text.

Graphic design by BGRAPHIC, Denmark

Printing, reproduction or quotation is authorised provided the source is acknowledged and a copy is forwarded to the ESPON EGTC in Luxembourg.

Contact: info@espon.eu

ESPON

EUROPEAN UNION

Co-financed by the European Regional Development Fund

Inspire Policy Making with Territorial Evidence

ESPON STUDY //

Geography of COVID-19 outbreak and first policy answers in European regions and cities

Annex : data collection process // December 2020

1 Summary of the methodology of data collection

As mortality data are not systematically produced at the NUTS2/NUTS3 regional level, it was necessary to create an original database compiling the statistics on the number of deaths produced in each State of the ESPON Space. This document serves as a guidance for collecting and updating the indicators used in this study (as described in the metadata file). It describes what has been done (step by step) and how to update the database continuously. Additional information is also given concerning data processing and cartography.

The steps are described below.

STEP 1: The general objective of the data collection was to find data at the finest scale (NUTS 2 or even NUTS 3). **Sources list: available at ESPON**

Most of the data were downloadable directly, but some data have been retrieved manually, for each day, for some of the sources. The goal was to successfully download data from official sources (Ministry of Health in particular).

For several countries, data were only available at NUTS 0 level (Greece, Bulgaria, Norway, Finland, Estonia, Latvia, Lithuania, Hungary and Slovenia). The data have been uploaded from the official website of the World Health Organisation https://worldhealthorg.shinyapps.io/covid/.

STEP 2: Once the raw data were found for each country, they were compiled into a single database. This involved creating a double-entry database (Date and NUTS ID) and then filling in the days for which we had values. When boxes remained empty (due to lack of daily data for example), the NULL values were replaced by the last previous value.

STEP 3: To produce maps showing weekly or monthly trends, a dynamic pivot table based on weekly or monthly values was created.

Databases are currently available for daily, weekly and monthly values. They were combined with the EUROPA GIS layer of the European NUTS to carry out spatial and cartographic studies. Two types of maps are available: (1) one polygonal map by region and (2) one smoothed map based on a spatial interaction model.

For the second one, the method chosen was the IDW. The weighted distance (IDW) is an interpolation of a point vector layer. The sample points are weighted during the interpolation so that the influence of one point relative to another decreases with the distance from the unknown point you want to create. Points are created from the NUTS centroids.

To create an animated map, values of the table per week (or month) must be selected and mapped weekly (or monthly). Then, the grid can be made with a website like "ezgif.com" (low quality) or with the software photoshop (good quality).

https://helpx.adobe.com/photoshop/how-to/make-animated-gif.html

2 Methodology step by step

Initial Step for the method that is described in this document: Create the following master folders: RAW_DATA, PROCESSED_DATA, BD AND RENDER.

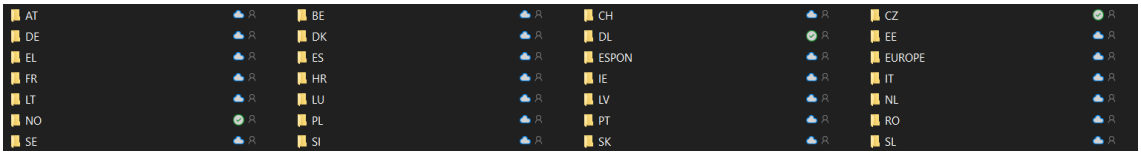
M : Methodology and data collection

P: Data processing

C : Data cartography

Step 1 (M): Create a "RAW_DATA" folder to store all download data (useful for the next steps).

Step 2 (M): Get the data: Download WHO data for all European countries (Level NUTS0). Store these data in the «RAW_DATA" > "WHO folder."



Step 3 (M): Create a file "metadata" to identify all sources used for the next step.

	FILE	TYPE	DESCRIPTION	THEMATIC	LASTUPDATE	COLUMN	DESCRIPTION OF THE COLUMN	TYPE C.
2	HOSPITALIZED_MONTHS_NEWIN	CSV	New hospitalizations per month (USED)	HOSPITALIZATION	28-oct	NUTS_ID	ID of the NUTS (0,1,2 or 3)	TEXT
3	HOSPITALIZED_MONTHS_NEWIN	CSV	New hospitalizations per month (USED)	HOSPITALIZATION	28-oct	1;2;3;...	Number of the month	NUMBER
4	HOSPITALIZED_WEEKS_NEWIN	CSV	HOSPITALIZED_WEEKS_NEWIN.csv: New hospitalizations per week (USED)	HOSPITALIZATION	27-oct	NUTS_ID	ID of the NUTS (0,1,2 or 3)	TEXT
5	HOSPITALIZED_WEEKS_NEWIN	CSV	HOSPITALIZED_WEEKS_NEWIN.csv: New hospitalizations per week (USED)	HOSPITALIZATION	27-oct	1;2;3;...	Number of the week	NUMBER
6	HOSPITALIZED_WEEKS	CSV	HOSPITALIZED_WEEKS.csv: Cumulative hospitalizations per week (USED)	HOSPITALIZATION	27-oct	NUTS_ID	ID of the NUTS (0,1,2 or 3)	TEXT
7	HOSPITALIZED_WEEKS	CSV	HOSPITALIZED_WEEKS.csv: Cumulative hospitalizations per week (USED)	HOSPITALIZATION	27-oct	1;2;3;...	Number of the week	NUMBER
8	HOSPITALIZED	CSV	Number of cumulative and no-cumulative hospitalizations per day and per NUTS (RAW)	HOSPITALIZATION	27-oct	DATE	Date of the data	DATE
9	HOSPITALIZED	CSV	Number of cumulative and no-cumulative hospitalizations per day and per NUTS (RAW)	HOSPITALIZATION	27-oct	DATE_NB	Date of the data on number	NUMBER
10	HOSPITALIZED	CSV	Number of cumulative and no-cumulative hospitalizations per day and per NUTS (RAW)	HOSPITALIZATION	27-oct	NUTS_ID	ID of the NUTS (0,1,2 or 3)	TEXT
11	HOSPITALIZED	CSV	Number of cumulative and no-cumulative hospitalizations per day and per NUTS (RAW)	HOSPITALIZATION	27-oct	NUTS_LEVEL	Level of the NUTS (0,1,2 or 3)	TEXT
12	HOSPITALIZED	CSV	Number of cumulative and no-cumulative hospitalizations per day and per NUTS (RAW)	HOSPITALIZATION	27-oct	HOSP_ALL	Cumulative hospitalized	NUMBER
13	HOSPITALIZED	CSV	Number of cumulative and no-cumulative hospitalizations per day and per NUTS (RAW)	HOSPITALIZATION	27-oct	HOSP_NEWIN	New hospitalized on the day	NUMBER
14	HOSPITALIZED	CSV	Number of cumulative and no-cumulative hospitalizations per day and per NUTS (RAW)	HOSPITALIZATION	27-oct	WEEK	Week of the day	NUMBER
15	DEATHS_MONTHS_NEWIN	CSV	DEATHS_MONTHS_NEWIN.csv: New deaths per month (USED)	DEATHS	28-oct	NUTS_ID	ID of the NUTS (0,1,2 or 3)	TEXT
16	DEATHS_MONTHS_NEWIN	CSV	DEATHS_MONTHS_NEWIN.csv: New deaths per month (USED)	DEATHS	28-oct	1;2;3;...	Number of the month	NUMBER

Step 4 (M): Get the data: Download chronological geographic datasets for each country from the sources identified.

Step 5 (M): Some sources cannot be downloaded directly in a single file; the download must be done file by file. This is the case of Italy.

Step 6 (M): Store data in the folders with the country's ISO code*, created within the master folder "RAW_DATA".

** The ISO codes for the countries are available at this link: https://www.nationsonline.org/oneworld/country_code_list.htm*

Step 7 (M): Create the "PROCESSED_DATA" folder. In this folder, the standard files will be stored. The aim of standardisation is to generate an identical table format for each country (see example below: column 1 = DATE; column 2 = DATE_NB, etc).

	A	B	C	D	E	F	G	H
1	DATE	DATE_NB	NAME	NUTS_ID	NUTS_LEVEL	DEATHS	WEEK	NEWIN_DEATHS
2	10/03/2020	43900	Région De Bru	BE1	NUTS1	1	11	1
3	11/03/2020	43901	Région De Bru	BE1	NUTS1	3	11	2
4	12/03/2020	43902	Région De Bru	BE1	NUTS1	4	11	1

Step 8 (M): Download the shapefiles for the European NUTS (Source: https://ec.europa.eu/eurostat/fr/web/gisco/geodata/reference-data/administrative-units-statistical-units/nuts); Use the shapefile on the scale "1:10 Million" over the year 2021.

Step 9 (P): All the data used in this study have been attached to NUTS 2021. If data has codes with NUTS 2018 or earlier, then correspondence table can be used to replace the old NUTS codes.

https://ec.europa.eu/eurostat/fr/web/nuts/history

Step 10 (P): Data processing country-by-country: The first step is to create a table (CSV or Excel) with the following table rows: date information, NUTS_ID or NAME_GEO, and associated values. The second step will be to join data with the NUTS layer to get the NUTS ID or NAME and its geographic level.

	A	B	C	D
1	DATE	DATE_NB	NAME	NUTS_ID
2	10/03/2020	43900	Région De Bru	BE1
3	11/03/2020	43901	Région De Bru	BE1
4	12/03/2020	43902	Région De Bru	BE1

Step 11 (P): Create the daily country file: Export the file with data concerning cases, deaths, or hospitalizations with the ID of the NUTS as a board in the "PROCESSED_DATA" master folder. This will be modified in the next step. The goal is to create a single file with all the dates per day and the NUTS.

Step 12 (P): Normalize the structure of the table (the table must include the date in this format :21/12/2020), the name of the NUTS, the NUTS ID, the geographic level of the NUTS (0, 1, 2, or 3), the number of "values" per day, the number of "values" accumulated per week, or per month, if necessary.
In order to facilitate the creation of pivot tables, it is recommended to generate a specific date format in Excel (DATEVALUE function) https://support.microsoft.com/en-us/office/datevalue-function-df8b07d4-7761-4a93-bc33-b7471bbff252

This is a preview example of the table format:

	A	B	C	D	E	F	G	H
1	DATE	DATE_NB	NAME	NUTS_ID	NUTS_LEVEL	DEATHS	WEEK	NEWIN_DEATHS
2	10/03/2020	43900	Région De Bru	BE1	NUTS1	1	11	1
3	11/03/2020	43901	Région De Bru	BE1	NUTS1	3	11	2
4	12/03/2020	43902	Région De Bru	BE1	NUTS1	4	11	1

Step 13 (P): Repeat step 12 for all countries, with CSV tables having the same structure. Save all tables in the «PROCESSED_DATA" master folder created in step 6.

Step 14 (P): Work Methodology: Create a "BDD" master folder to place the data that will be used to make the maps (optional it is one method to work).

Step 15 (M): Group all data from the "PROCESSED_DATA" folder into a single XLS file. At the time when this database was created, it had about 150,000 lines.

Step 16 (P): With a dynamic pivot table, create a table with ID_NUTS (in rows) and WEEK (in columns) with the maximum value for the week. Copy this table to another tab.

Step 17 (P): Fill in the data on weekly data that are empty by the previously available value. This step will allow us not to have empty areas on a future animation. Do the same with the monthly data if they are available.

Step 18 (M): Save this table in the "BDD" master table.

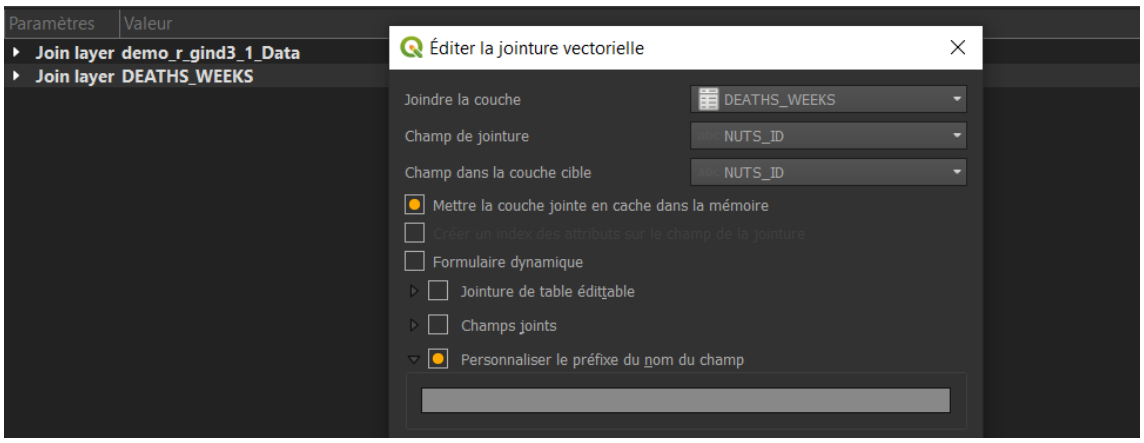
Step 19 (M): Download the NUTS population from the Eurostat database available online. https://ec.europa.eu/eurostat/fr/data/database.

Step 20 (P): Join the weekly values file of the population associated with NUTS. To do this, you can either make a join from GIS software on the NUTS ID or make a Ctrl+F on Excel to search for items in a table or range per line. To learn how to do: see for example: https://youtu.be/82aRDdJdhI0

Step 21 (P): Calculate the number of "values" per 100,000 inhabitants. The formula used is: Number of deaths per NUTS * 100000 / Number of inhabitants in the NUTS.

⇨ At this point, a new raw will be created in the table representing the density of "values" per week and by NUTS.

Step 22 (P): Join the geographic layer and the weekly densities table to have data per week with geographic entities.



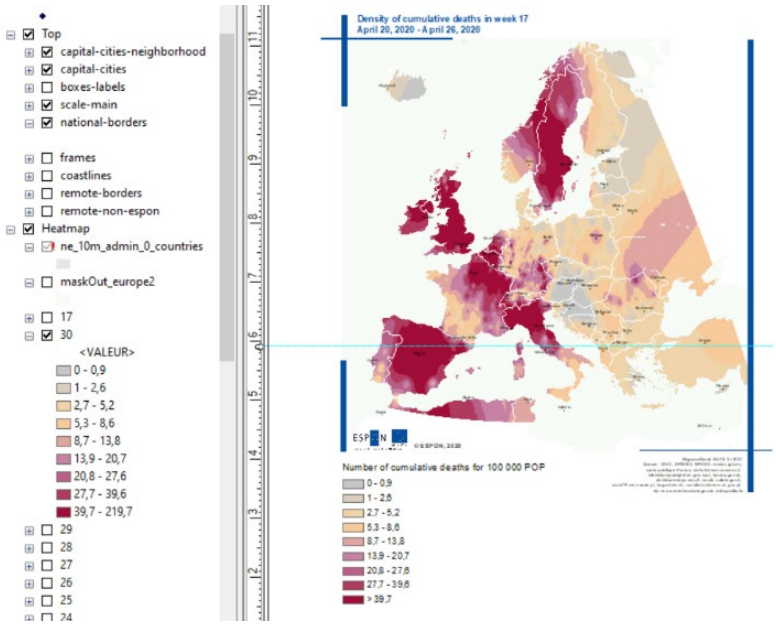
Step 23 (C): Classify the rendering by going into the layer settings to have geographic features with the highest-level nuts available. The goal is to keep the finest NUTS level in a given territory.

Step 24 (C): Create maps for depicting each week trends using polygons. To have the same scale on the maps, a scattering of the week's most "valuable" data will allow all other weeks' style and semiology to be adapted.

Step 25 (P): Create polygon centroids (points) and select NUTS with the highest level by country.

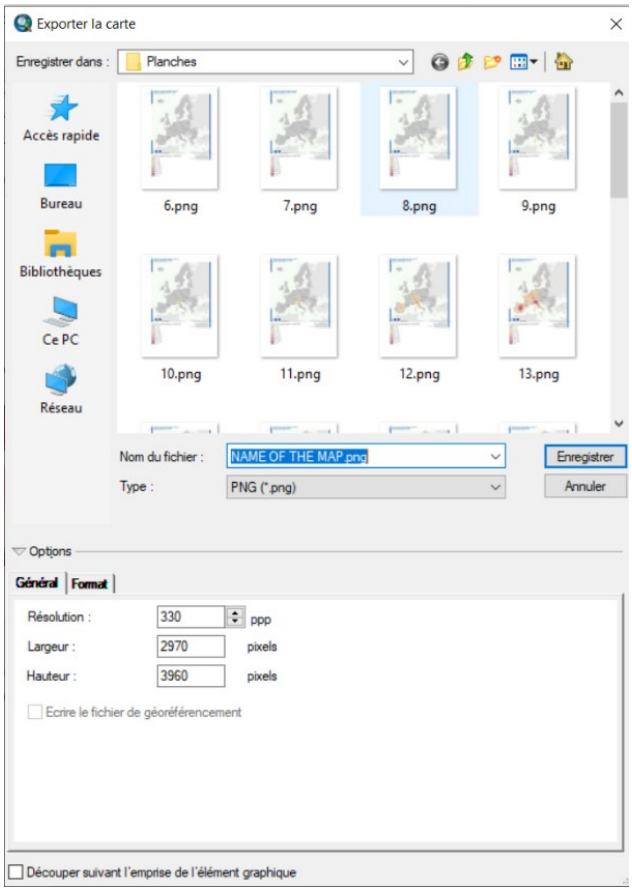
Step 26 (C): With the points generated, create an IDW spatial interaction model. ArcGIS is the software used. Distance Weighted (IDW) is an interpolation of a layer of point vectors. In the IDW interpolation method, the sample points are weighted during interpolation, such that the influence of one point relative to another decline with distance from the unknown point you want to create. This method can be found in the spatial analyst extension in ArcGIS Software. https://desktop.arcgis.com/fr/arcmap/10.3/tools/3d-analyst-toolbox/how-idw-works.htm

Step 26 (C): Create a unique discretization that will be applied with the IDW model for each week.

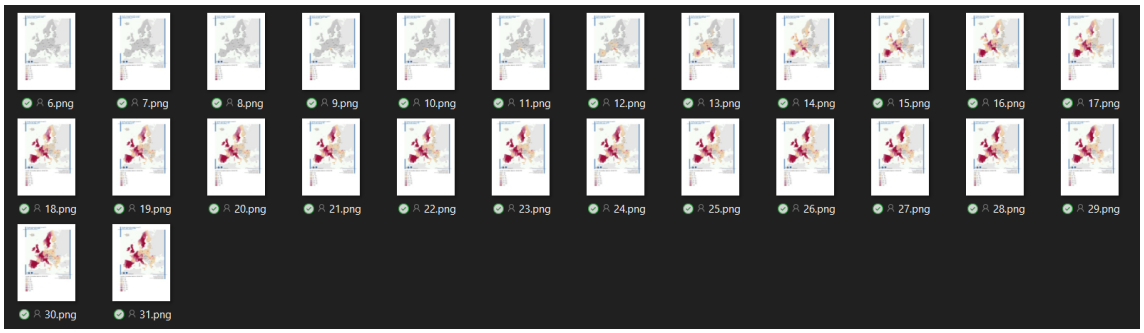


Step 27 (M): Create a “RENDER” master folder. It contains all the documents (such as maps, sources, etc.) that are definitive.

Step 28 (C): Export maps in a "MAP_EXPORT" folder in “RENDER". Export maps with a minimum resolution of 300 IDPI.



Step 29 (C): To create a gif, it is possible to use a website like "ezgif.com" but the quality will be low or passable or use Photoshop or other software to get the same resolution as the exported boards.



3 Sources

- BE
https://epistat.wiv-isp.be/covid/

- DE
https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/dd4580c810204019a7b8eb3e0b329dd6_0/data
https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/917fc37a709542548cc3be077a786c17_0
https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Fallzahlen.html

- DK
https://www.ssi.dk/sygdomme-beredskab-og-forskning/sygdomsovervaagning/c/covid19-overvaagning
https://www.sst.dk/da/corona/tal-og-overvaagning

- ES
https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov-China/documentos/Actualizacion_36_COVID-19.pdf

- FR
https://www.data.gouv.fr/fr/datasets/donnees-hospitalieres-relatives-a-lepidemie-de-covid-19/

- IT
https://statistichecoronavirus.it/crea-dashboard/
http://www.salute.gov.it/portale/nuovocoronavirus/dettaglioContenutiNuovoCoronavirus.jsp?lingua=english&id=5367&area=nuovoCoronavirus&menu=vuoto

- NL
https://www.rivm.nl/coronavirus-covid-19/actueel/wekelijkse-update-epidemiologische-situatie-covid-19-in-nederland
https://www.arcgis.com/apps/opsdashboard/index.html#/cfc2084c995c40e7ae72254029bf6251

- PT
https://covid19.min-saude.pt/ponto-de-situacao-atual-em-portugal/
https://github.com/dssg-pt/covid19pt-data

- CH
https://github.com/openZH/covid_19
/https://www.bag.admin.ch/bag/en/home/krankheiten/ausbrueche-epidemien-pandemien/aktuelle-ausbrueche-epidemien/novel-cov/situation-schweiz-und-international.html

- CZ
https://covid19.who.int/
https://onemocneni-aktualne.mzcr.cz/api/v2/covid-19

- HR
https://www.koronavirus.hr/podaci/otvoreni-strojno-citljivi-podaci/526
https://www.koronavirus.hr/latest-news/covid-19-weekly-report-august-31st/786

- AT

Data collected with a need for authorization (have to wait one month between the request and the obtention of the excel file): contact:

Dr. Michael Hummer michael.hummer@goeg.at
Gesundheit Österreich GmbH
Stubenring 6
1010 Wien
T: +43 1 515 61-138
M: +43 676 84 81 91 610

https://www.sozialministerium.at/Informationen-zum-Coronavirus/Neuartiges-Coronavirus-(2019-nCov).html
https://experience.arcgis.com/experience/fb603473e1f74f0bbae48155ff238565/page/page_2/

- HU

https://covid19.who.int/

- PL

ESPON contact Poland: Katarzyna.Powierza@mfipr.gov.pl
https://www.gov.pl/web/koronawirus/wykaz-zarazen-koronawirusem-sars-cov-2
https://www.arcgis.com/apps/opsdashboard/index.html#/deaceebc69a3412c8b7699e3e025e213
http://www.informacjakryzysowa.pl/en/aktualnosci/covid-19-na-mapach-polska

- SK

https://korona.gov.sk/koronavirus-na-slovensku-v-cislach/
https://korona.gov.sk/en/coronavirus-covid-19-in-the-slovak-republic-in-numbers/

- SI

https://covid-19.sledilnik.org/sl/stats
https://www.nijz.si/en

- CY

https://covid19.who.int/
https://covid19.ucy.ac.cy/

- EL/GR

https://covid19.who.int/
https://covid19.gov.gr/covid19-live-analytics/

- LT

http://atviriduomenys.nvsc.lt/
https://covid19.who.int

- LU

https://covid19.who.int/
https://data.public.lu/en/datasets/donnees-covid19/

- MT

https://covid19.who.int/

- RO

https://coronavirus-esriro.hub.arcgis.com/
https://esriro.maps.arcgis.com/apps/opsdashboard/index.html#/ec6ee9017ee549e992e96f1b8bd7d080

- EE

https://covid19.who.int/EE

- FI

https://covid19.who.int/FI

- IE

https://covid19.who.int/LV
https://covid19ireland-geohive.hub.arcgis.com/datasets/c8208a0a8ff04a45b2922ae69e9b2206_0

- SE

https://www.folkhalsomyndigheten.se/
https://ago-item-storage.s3.us-east-1.amazonaws.com/

- LI
https://covid19.who.int/
https://github.com/openZH/covid_19/tree/master/fallzahlen_kanton_total_csv

- NO
https://covid19.who.int/

- UK
https://www.gov.uk/government/news/weekly-covid-19-surveillance-report-published
https://coronavirus.data.gov.uk/details/deaths

Co-financed by the European Regional Development Fund

Inspire Policy Making with Territorial Evidence

ESPON 2020

ESPON EGTC
4 rue Erasme, L-1468 Luxembourg
Grand Duchy of Luxembourg
Phone: +352 20 600 280
Email: info@espon.eu
www.espon.eu

The ESPON EGTC is the Single Beneficiary of the ESPON 2020 Cooperation Programme. The Single Operation within the programme is implemented by the ESPON EGTC and co-financed by the European Regional Development Fund, the EU Member States and the Partner States, Iceland, Liechtenstein, Norway and Switzerland.

Disclaimer
This delivery does not necessarily reflect the opinion of the members of the ESPON 2020 Monitoring Committee.