

# AI4COVID: Artificial intelligence and geographical information for monitoring and prediction of Covid-19 outbreak

Jamal Jokar Arsanjani  
Geoinformatics Research Group  
Aalborg University Copenhagen



[View Scopus Profile](#)

## Jamal Jokar Arsanjani

Cartography and Geoinformation  
Professor, Department of Planning  
Professor, The Technical Faculty of IT and Design  
Professor, Land Management And Geoinformatics  
 <https://orcid.org/0000-0001-6347-2935>

Phone  
9356 2323

Email  
[jja@plan.aau.dk](mailto:jja@plan.aau.dk)



## Frederik Seerup Hass

Research assistant, Department of Planning  
!!Videnskabelig assistent, Department of Planning  
Research assistant, The Technical Faculty of IT and Design  
!!Videnskabelig assistent, The Technical Faculty of IT and Design  
Research assistant, Land Management And Geoinformatics  
!!Videnskabelig assistent, Land Management And Geoinformatics

Email  
[frederiksh@plan.aau.dk](mailto:frederiksh@plan.aau.dk)

# Objectives

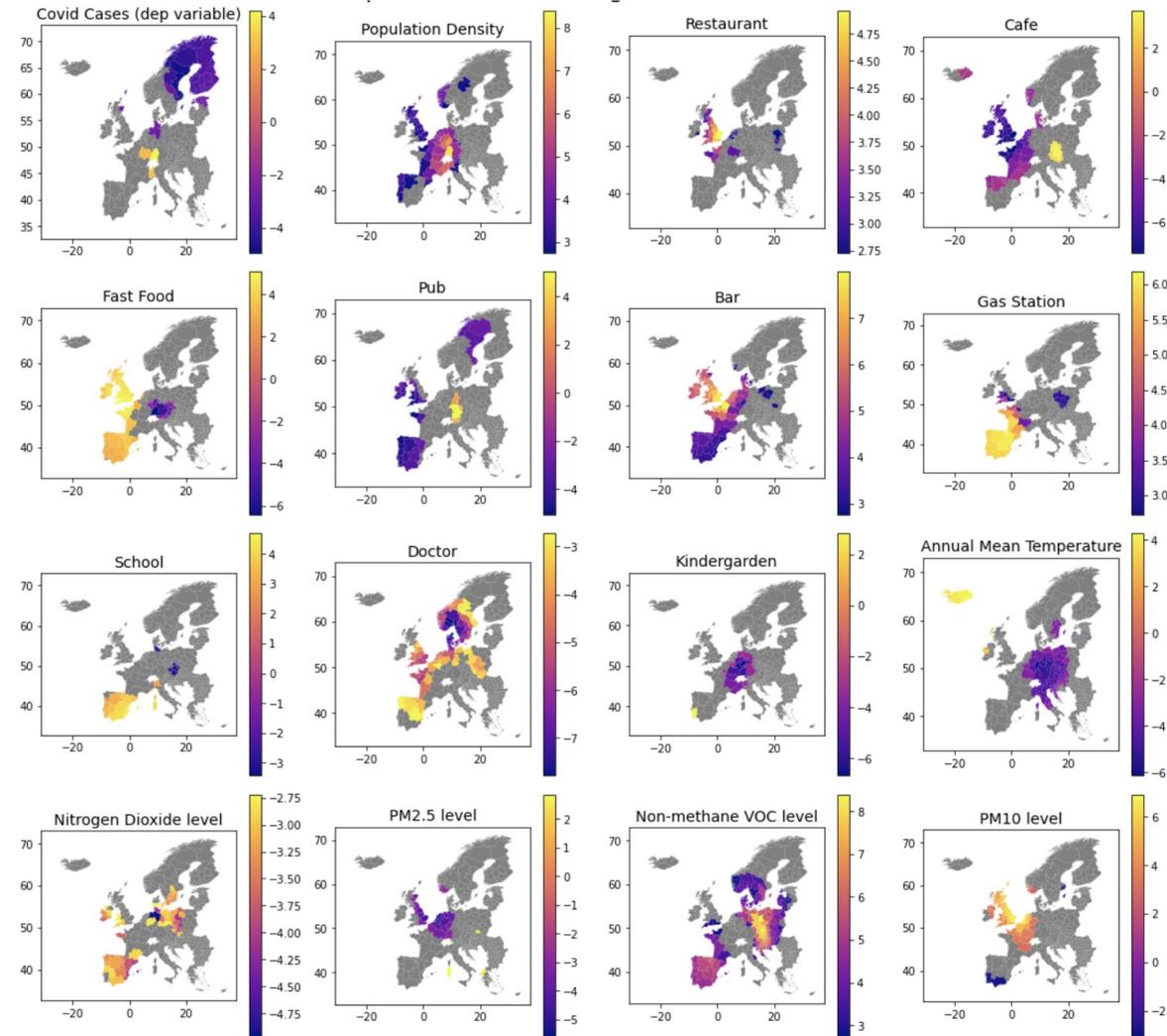
Our study aimed at

- *presenting a data-driven approach for exploring the spatio-temporal patterns of the pandemic over a regional scale, i.e., Europe and a country scale, i.e., Denmark,*
- *Exploring what geographical variables potentially contribute to expediting the spread of the pandemic so that they could be used for lockdown/re-opening planning.*

# Data

We used

- **Official regional infection rates (source: European Centre for Disease Prevention and Control)**
- Population Density,
- Public places per capita e.g., Restaurants, Cafes, Fast Food Places, Pubs, Bars, Gas Stations, Schools, Medical Offices, Kinder gardens,
- Annual mean temperature,
- Nitrogen dioxide,
- Particulate matter < 2.5  $\mu\text{m}$  (PM2.5),
- Non-methane VOCs,
- Particulate matter < 10  $\mu\text{m}$  (PM10)



# Methods-1

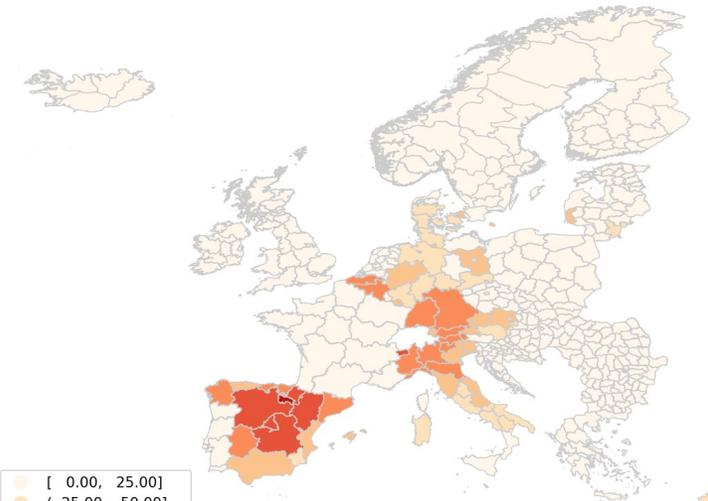
- Historical development of the pandemic

2020-03-01

Covid-19 case numbers over the past two weeks



2020-Week14



Covid-19 cases per 100.00 inhabitants over the past two weeks

The map is created by Frederik Hass under supervision of Jamal Jokar Arsanjani as part of a research project at Department of Planning, Aalborg University called "AI4Covid: Artificial Intelligence for Covid-19 analysis" funded by the European Open Science Cloud.

2020-03-01



Covid-19 cases per 100.00 inhabitants over the past 14 days

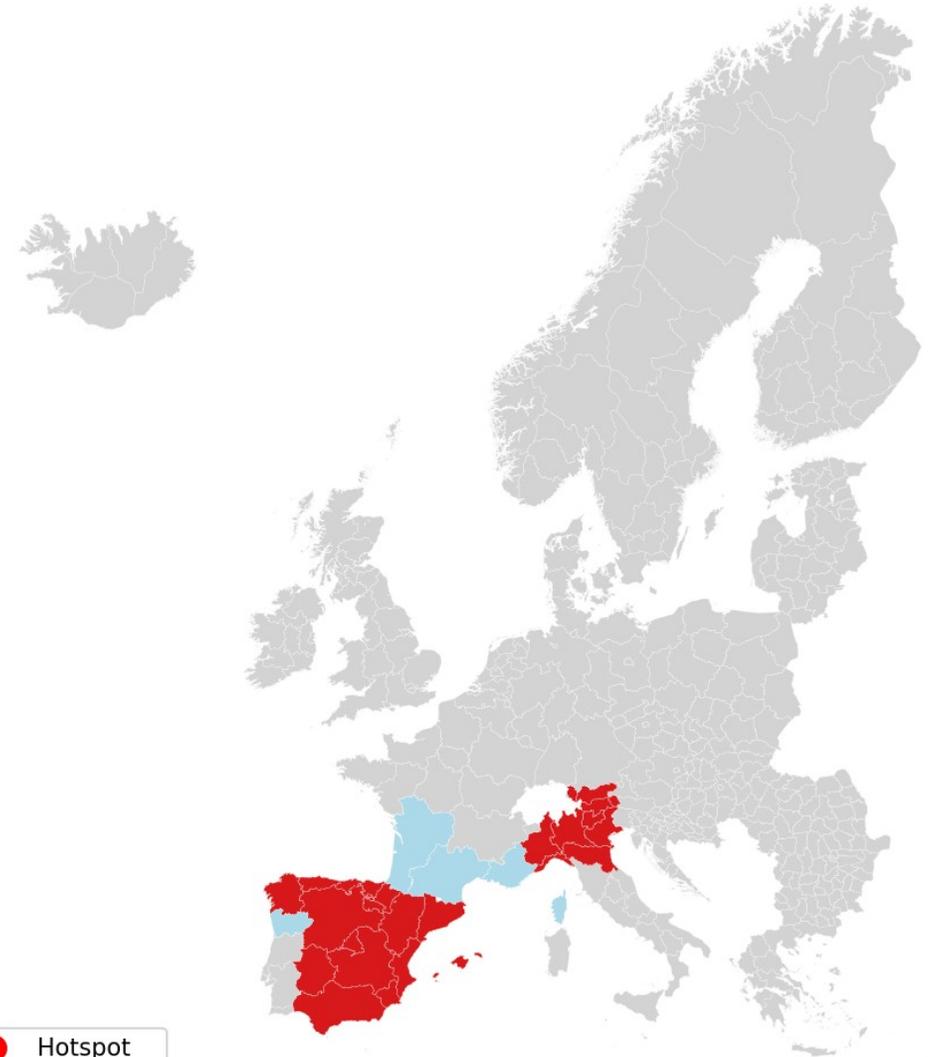
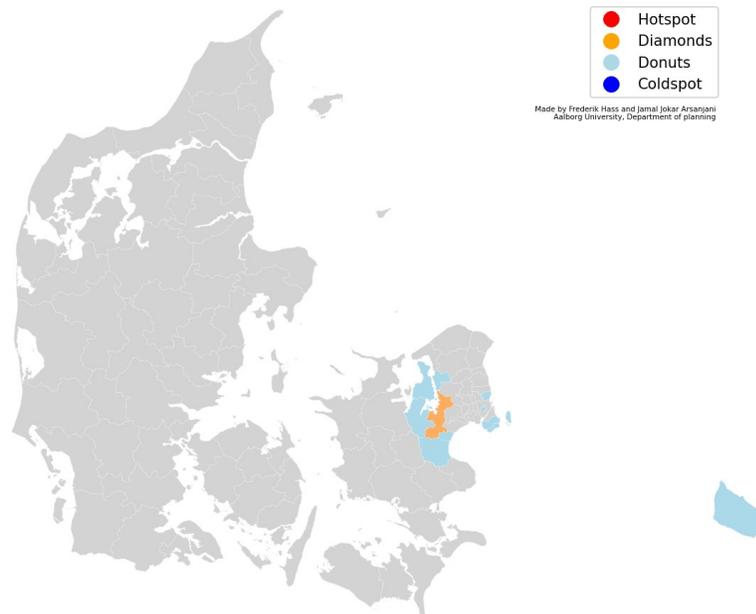
The map is created by Frederik Hass under supervision of Jamal Jokar Arsanjani as part of a research project at Department of Planning, Aalborg University called "AI4Covid: Artificial Intelligence for Covid-19 analysis" funded by the European Open Science Cloud.

# Methods-2

- we measured **spatial autocorrelation** to extract relevant indicators that could explain the dynamics of the pandemic.

2020-03-01

Covid-19 14-day infection rates  
Global Spatial Autocorrelation

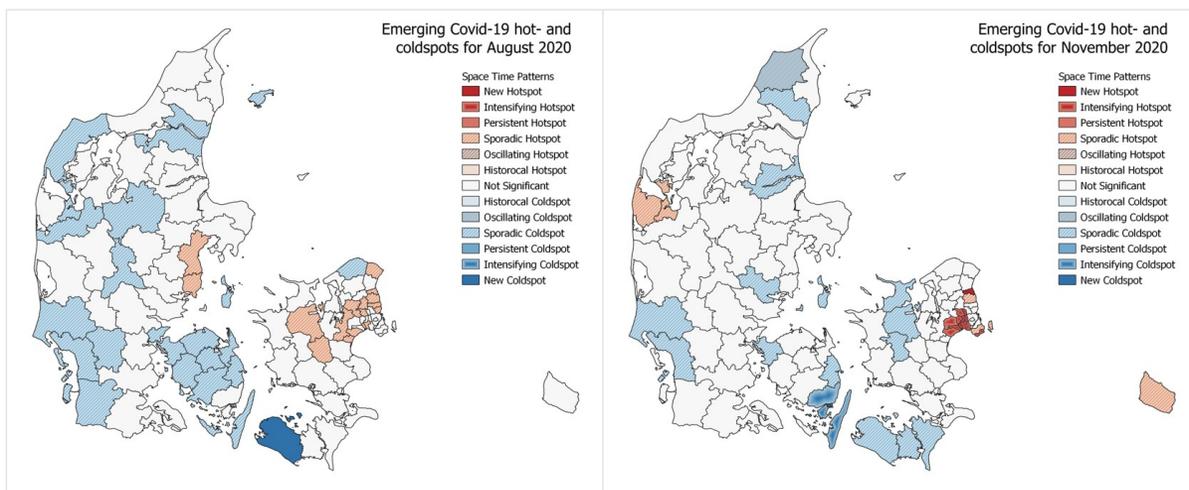


Weekly Covid-19 Spatial Autocorrelation

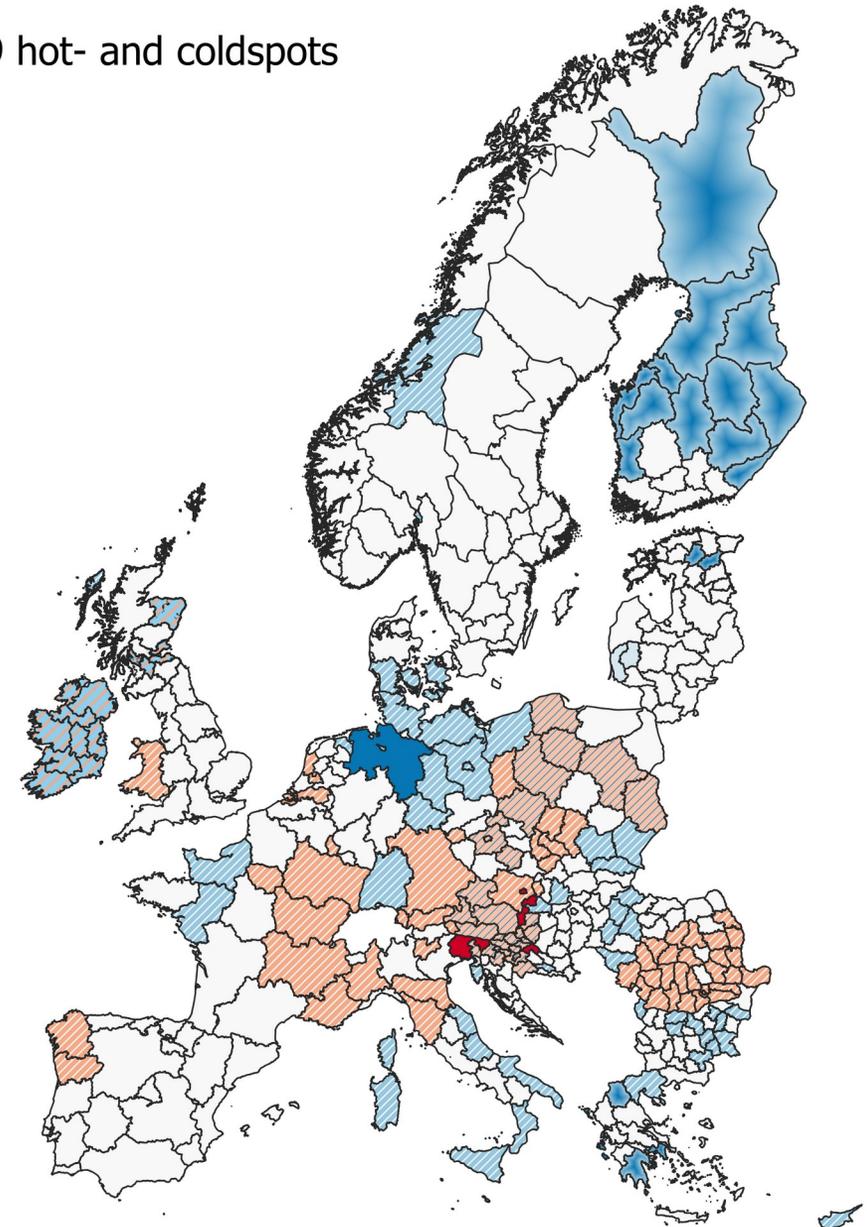
Hot- and coldspots indicates clusters of high and low infection rates.  
Donuts are regions with low infection-rates surrounded by areas with high infection-rates.  
Diamonds are regions with high infection-rates surrounded by regions with low infection-rates

# Methods-3

- we measured **space-time autocorrelation** to extract relevant indicators that could explain the dynamics of the pandemic.



Emerging Covid-19 hot- and coldspots



# Methods-4

- In order to explore the potential correlation between the chosen underlying factors and the pandemic spread, we applied:
  - statistical methods e.g., ordinary least squares (OLS), geographically weighted regression (GWR)

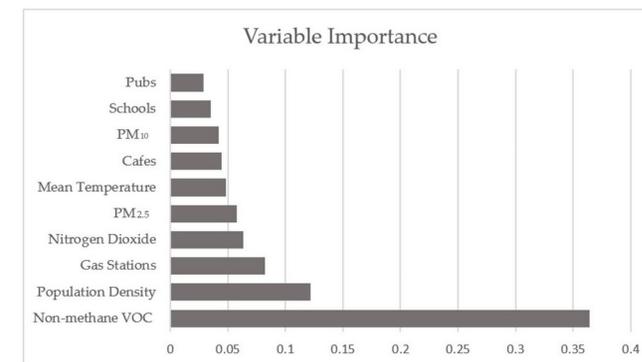
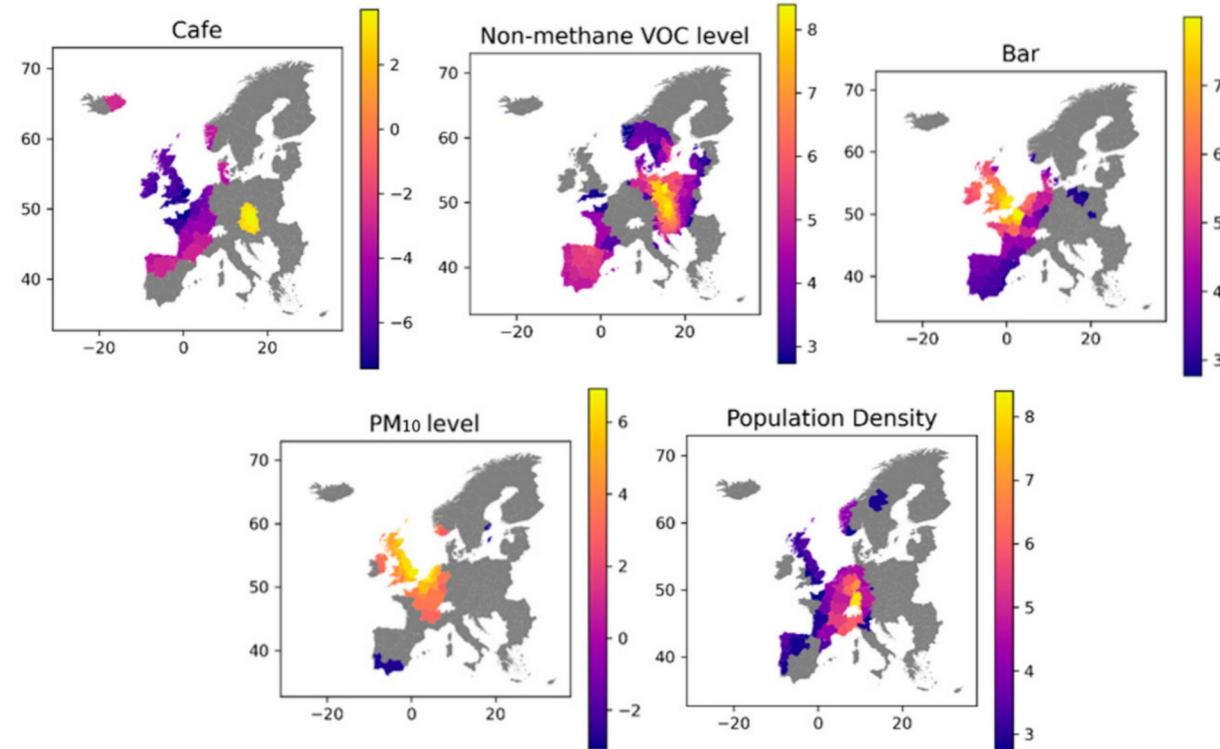
Model	Res Sum Squares	AIC	AICc	R <sup>2</sup>	Adjusted R <sup>2</sup>
OLS	219.230	927.850	931.448	0.453	0.432
GWR	66.922	654.585	752.719	0.833	0.765

- machine learning methods

Area	Nr. of Regions	Model	Accuracy	MAE
Europe	401	Random Forest	52%	0.079
		Lasso	36%	0.153
		Support Vector Regression	51%	0.151
Europe—Hot- & Coldspots	128	Random Forest	76%	0.099
		Lasso	74%	0.285
		Support Vector Regression	73%	0.256
Europe—Hotspots	56	Random Forest	19%	0.177
		Lasso	- %	0.217
		Support Vector Regression	- %	0.221
Denmark	98	Random Forest	61%	0.095
		Lasso	35%	0.233
		Support Vector Regression	41%	0.232

# Findings

- Our findings indicate that **population density, public places** such as cafes and bars, and **pollution levels** are the most influential explanatory variables.



# Lessons learned

- Pollution levels can be explicitly used to monitor lockdown measures and infection rates.
- Through identifying hot- and cold spots of infections, varying lockdown measures can be reinforced.
- The most influential driving forces could be used for prediction of upcoming infections.
- Some soft factors could not be incorporated in our models since they are not measurable.
- There is a need for fine scale data e.g., neighbourhood level as opposed to country/provincial/municipality level.

# Conclusions

- Geographical data, tools and methods can provide unique insights to understanding the evolution of the pandemic.
- Our choice of data and methods along with the achieved results can empower health authorities and decision makers with an interactive decision support tool, which can be useful for imposing geographically varying lockdowns/re-openings.
- As per future work, we suggest including further variables that reflect lockdown levels of individual regions over time. Such a variable could be made from quantification of adopted policies and pollution levels.

# Thank you for your attention

Hass, F.S.; Jokar Arsanjani, J. The Geography of the Covid-19 Pandemic: A Data-Driven Approach to Exploring Geographical Driving Forces. *Int. J. Environmental Research and Public Health* 2021, *18*, 2803. <https://doi.org/10.3390/ijerph18062803>

**Our data and approach** is accessible at [https://github.com/jimjoker/Covid-19\\_geographic\\_analysis](https://github.com/jimjoker/Covid-19_geographic_analysis)

Questions and comments should be directed to:

[jja@plan.aau.dk](mailto:jja@plan.aau.dk)

Tweeter: @jamaljokar