

# A statistical analysis on factors that contributed/slowed down the spread of COVID-19

Project for the exam: Machine learning, statistical learning, deep learning and artificial intelligence - Unsupervised Learning



**FINAL GOAL:** Survey the statistical correlations between the COVID-19 cases/deaths and a selected set of attributes for different clusters via the principal component analysis

**CLUSTERS:** Provinces (IT), Regions (IT), Countries

# Cluster features

## Provinces (IT)

- ▶ Unemployment '19
- ▶ Private Transport '12
- ▶ Air Quality '19
- ▶ Public Transport '12
- ▶ Density '19
- ▶ Mean income '19

## Regions (IT)

- ▶ Mortality
- ▶ Pop. for GP
- ▶ Mean Income '19
- ▶ Number of Visits for Pop. '17
- ▶ Public Transport
- ▶ Density '19
- ▶ Mean income '19
- ▶ Tests
- ▶ LEA '17
- ▶ Number of visit '17
- ▶ Public structures '17

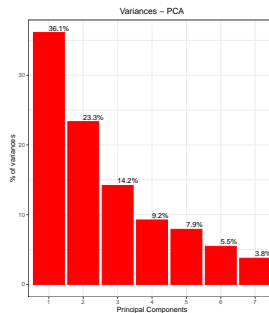
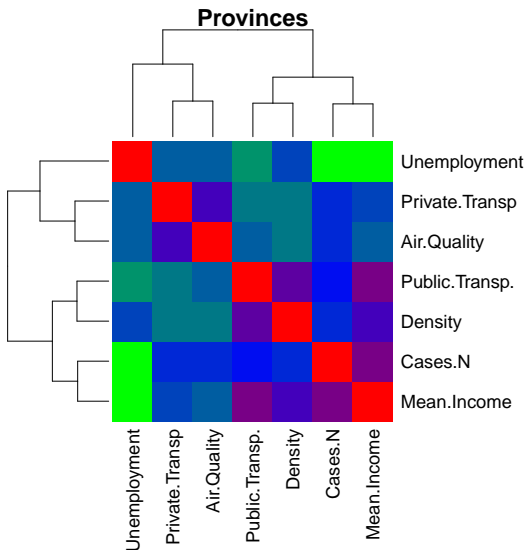
## Countries

- ▶ PM<sub>2.5</sub> '16
- ▶ Traffic Mortality '16
- ▶ Pollution Mortality '16
- ▶ GDP pro-capita '19
- ▶ Health expenditure '17
- ▶ UHC '17

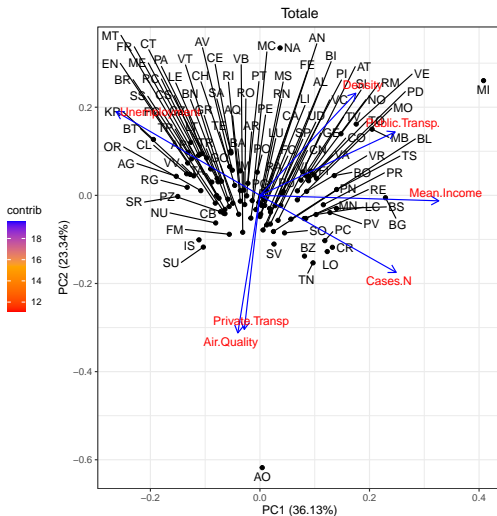
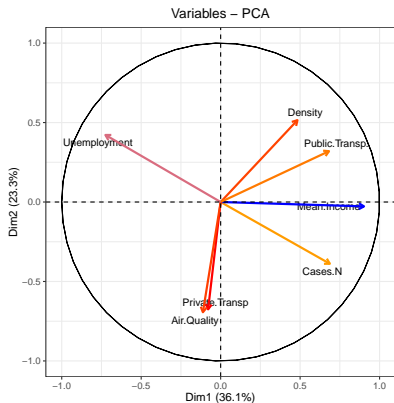
**TEMPORAL INTERVAL:** Provinces up to 24/08/2020; Regions up to 24/08/2020; Countries up to 27/08/2020

**SOURCES:** Protezione Civile, Istituto Italiano di Statistica, Ministro delle Finanze, Ministro della Salute, World Health Organization and World Bank

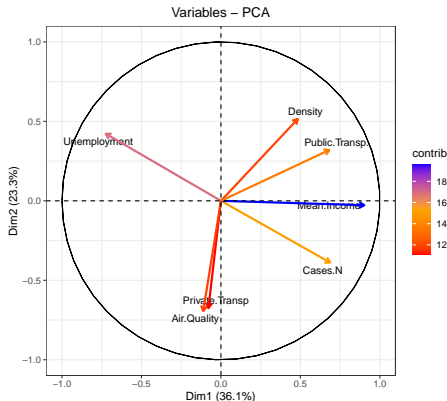
# Provinces



# Provinces

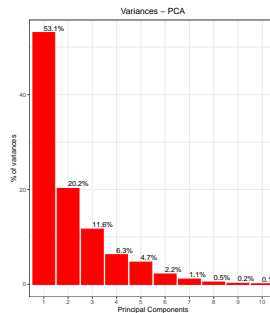
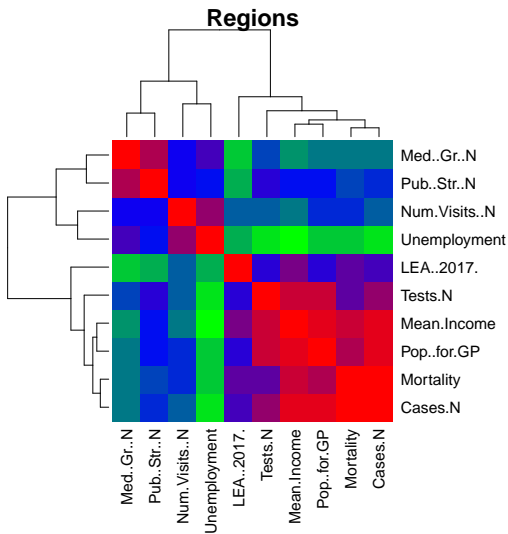


# Provinces - Conclusions



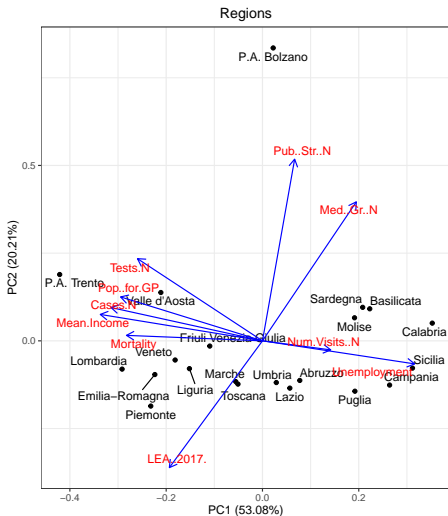
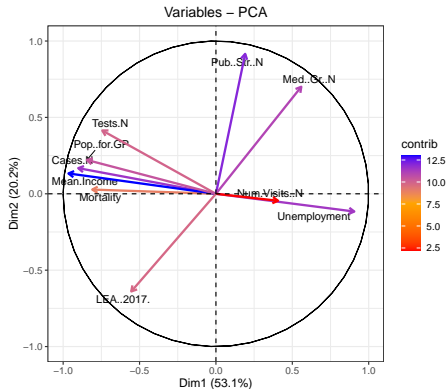
- ▶ The public transport, the density, the normalized cumulative cases and the mean income are positively correlated
- ▶ A higher density, a higher public transport demand and a higher income increase the rate of contact between the individuals [3, 2]
- ▶ Rich people can spend more money for social events or perhaps in to travels thus they increase their connectivity (the opposite happens for the unemployment)[4]

# Regions

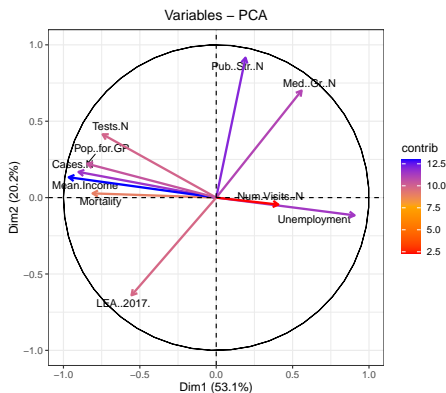




# Regions

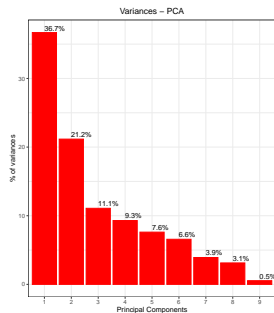
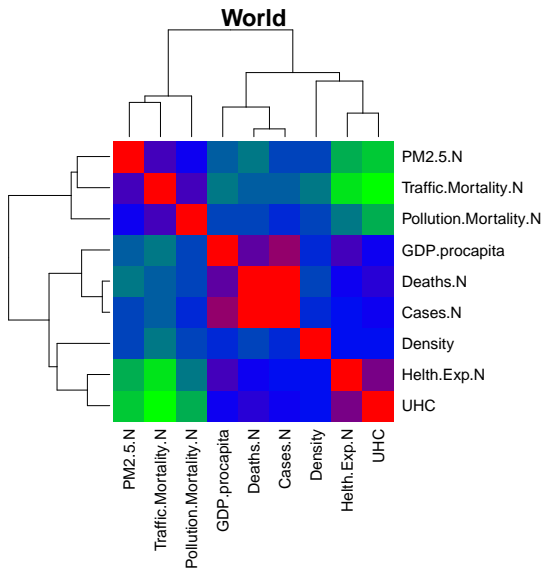


# Regions - Conclusions

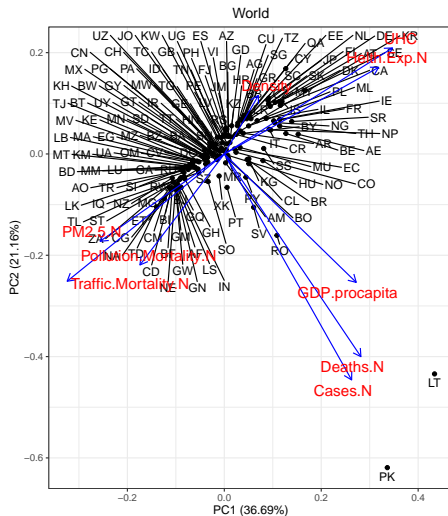
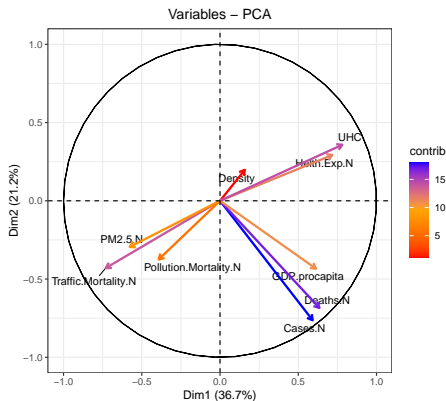


- ▶ Positive correlation between the number of people for general practitioner(GP) and the number of cases
- ▶ A higher number of normalized visit is negatively correlated with the number of cases and the mortality
- ▶ The GP play the role of sensor for the COVID-19 cases:if their number and their presence is reduced, the health service is less sensitive to find new cases and the diffusion of the virus is favoured [1]

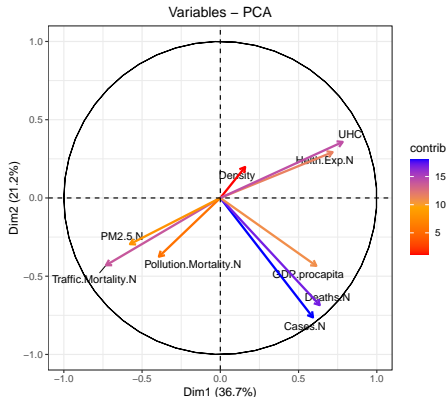
# Countries



# Countries



# Countries - Conclusions



- ▶ The GDP pro-capita seems to be statistically correlated with the cases and the deaths
- ▶ A possible objection to this link can be that only the richest countries can invest money for the tests for the validation of a COVID19 case: however it can be noted from plots that the normalized health expenditure as well the UHC index are basically orthogonal with respect to the normalized COVID19 cases and deaths

# References I

- [1] Roberto Galullo and Angelo Mincuzzi.  
<https://24plus.ilsole24ore.com/art/ADMIadK>.
- [2] World Health Organization. Centre for Health Development and World Health Organization. *Hidden cities: unmasking and overcoming health inequities in urban settings*. World Health Organization, 2010.
- [3] Carl-Johan Neiderud. “How urbanization affects the epidemiology of emerging infectious diseases”. In: *Infection ecology & epidemiology* 5.1 (2015), p. 27060.
- [4] Simone Weyers et al. “Low socio-economic position is associated with poor social networks and social support: results from the Heinz Nixdorf Recall Study”. In: *International Journal for Equity in Health* 7.1 (2008), p. 13.