# Description of data for the comparison of COVID-19 mortality in European and North American geographic entities 

M. Ribaud ${ }^{1}$, D. Martinetti ${ }^{1}$, and S. Soubeyrand ${ }^{1}$<br>${ }^{1}$ INRAE, BioSP, 84914 Avignon, France

May 18, 2021

## 1 References

Mélina Ribaud, Edith Gabriel, Joseph Hughes, Samuel Soubeyrand. Identifying potential significant factors impacting zero-inflated proportions data. 2021. <hal-02936779v3> https://hal.archives-ouvertes.fr/hal-02936779v3

Soubeyrand S., Ribaud M., Baudrot V., Allard D., Pommeret D., Roques L. (2020). COVID-19 mortality dynamics: The future modelled as a (mixture of) past(s). PLoS ONE 15(9): e0238410. https://doi.org/10.1371/journal.pone. 0238410

## 2 Format of data and description of variables

Data are provided in two tab-delimited text files. The rawFactors.txt file contains raw factors for all the geographic entities (countries, states and provinces; identified by the geographic_entity column) considered by Ribaud et al. (2021). The diffFactors.txt file contains for each source-receptor pair considered by Ribaud et al. (2021), the probability that the receptor (geographic_entity_receptor column) follows the mortality dynamics of the source (geographic_entity_source column) in column proba (see Soubeyrand et.al., 2020 fro the computation of this probability) as well as differences between the respective raw factors. The factors and the difference between factors in rawFactors.txt and diffFactors.txt, respectively, are listed in the following table.

| Category | Variable | Description | Unit |
| :---: | :---: | :---: | :---: |
| Economy | gdp2019 | Gross domestic product in 2019 | M ${ }^{\text {8 }}$ |
|  | gdp_capita | Gross domestic product per capita in 2019 | \$ |
|  | healthexp | Health expenditure | M ${ }^{\text {S }}$ |
| Demography | pop | Total population | units |
|  | density | Population density | units per $\mathrm{km}^{2}$ |
|  | urbanpop | Percentage of population living in urban areas |  |
|  | pop_female | Percentage of female | \% |
|  | pop_male | Percentage of male | \% |
|  | pop_tot_0_14 | Percentage of the total population in the age group 0-14 | \% |
|  | pop_tot_15_64 | Percentage of the total population in the age group 15-64 | \% |
|  | pop_tot_65_up | Percentage of the total population in the age group 65 or more | \% |
|  | pop_female_0_14 | Percentage of the female population in the age group 0-14 | \% |
|  | pop_female_15_64 | Percentage of the female population in the age group 15-64 | \% |
|  | pop_female_65_up | Percentage of the female population in the age group 65 or more | \% |
|  | pop_male_0_14 | Percentage of the male population in the age group 0-14 | \% |
|  | pop_male_15_64 | Percentage of the male population in the age group 15-64 | \% |
|  | pop_male_65_up | Percentage of the male population in the age group 65 or more | \% |
|  | mediange | Median age | years |
|  | life_expectancy | Life expectancy at birth | years |
| Health | lung | Death rate for lung diseases per 100,000 people | units |
|  | fertility | Average number of children per woman | units |
|  | obesity | Percentage of obese people within the popoulation | \% |
|  | smokers | Percentage of smokers within the population | \% |
| Healthcare System | hospibed | Number of hospital beds per 1,000 people | units |
|  | physicians_per_1K | Number of physicians per 1,000 people | units |
|  | nurses_per_1K | Number of nurses per 1,000 people | units |
| Climate | tmin | Average minimum temperature in the first semester | ${ }^{\circ} \mathrm{C}$ |
|  | $t$ max | Average maximum temperature in the first semester | ${ }^{\circ} \mathrm{C}$ |
|  | prec | Average precipitation in the first semester | mm |
|  | avghumidity | Average relative humidity | \% |

