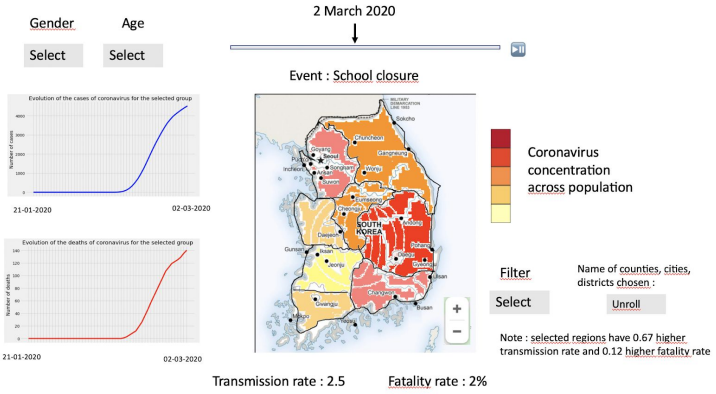
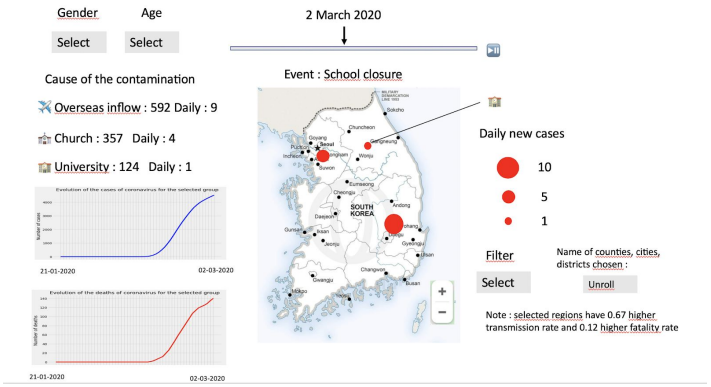
# Milestone 2

Data Visualization Project 2020, Team SolEckBis, Lucas Eckes, Anton Soldatenkov, Frédéric Bischoff

In this document we will present the visualizations we want to provide according to our problematic defined in Milestone 1. For more clarity, we split the description of the graphs in this document in three main parts: **1.** Show the ***spreading*** of the disease in the different regions of South Korea, **2.** See if some ***characteristics of a city*** play a role in the propagation (prop. of persons in the elderly, nb of schools ...), **3.** See if the ***weather*** fosters to propagate the weather (wind, temp., humid...).

## I. Spreading of Covid-19 in the South Korean population



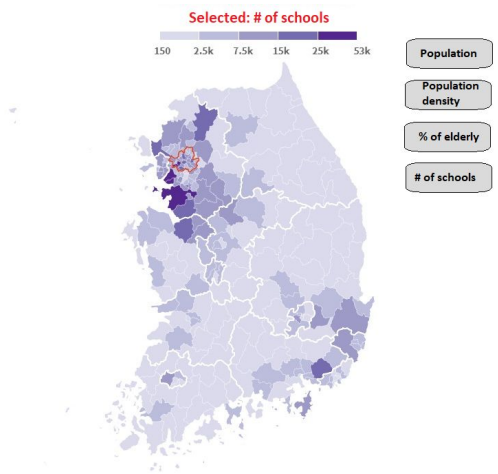

The first section will involve the timewise variation of the Covid-19 pandemic in South Korea. From the topojson data, a map will be created where the color will indicate the concentration of cases in each region. The color palette will be selected via color brewer or I want hue. The administrative divisions of South Korea include provinces, cities, counties and districts. We have information from most of the possible levels. A zooming in will be useful for districts and counties numbers whereas a zooming out will allow to visualize provinces or cities.

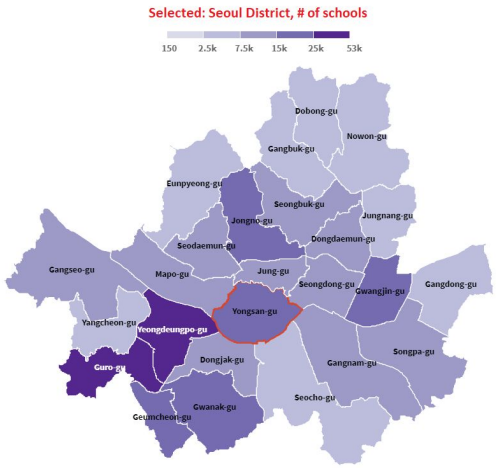
Furthermore a play and pause button will allow to select or to scroll the dates. Special events will be announced (special immigration procedure from China, Shincheonji Church event, ...). A curve indicating the total number of cases and deaths, the transmission rate and fatality rate will be displayed. We will be able to select the gender or the age group and the map and curves will adapt automatically. We then will be able to filter regions from their densities, their elderly population ratio or the number of elementary schools. The names of the selected regions will be displayed. In a greater extent, it will be possible to erase or write a region in the selected regions case.

In the next part, the visualization will be centered more on the daily cases and their causes. The structure will be similar but now circles of different sizes will represent the number of daily cases. It will indicate the causes of the contamination by a symbol. On the left side, the causes of contamination and their total numbers at this date will be presented. To develop even further our ideas other color circles may represent the daily deaths number.

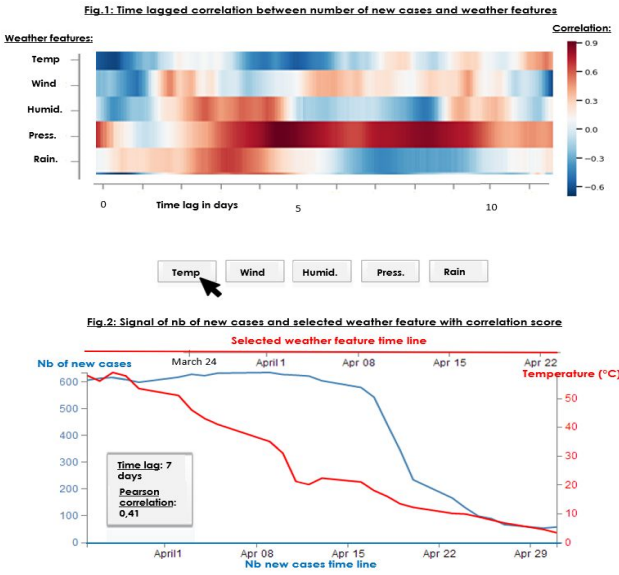
## II. Role of city morphology

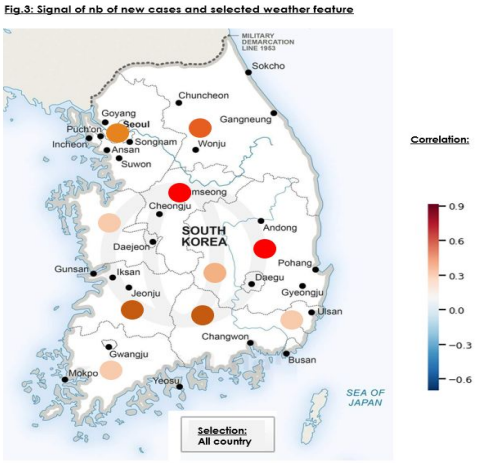
Here we want to investigate the dependence (or independence) of the epidemic spreading on different characteristic of cities and provinces, such as population density, percentage of elderly people, number of schools, churches, shopping centres and other socio-cultural objects per km2. The map will be a choropleth without timeline (we have not decided on the color scale yet) with several buttons that will change the displayed indicator. An option to click on the region to get an enlarged view and more detailed information will also be available, the scale will be adapted automatically. We will also consider doing correlation plots between city morphology and infection rates.




# III. Impact of the weather




We want to provide graphs that allows to explore any potential correlations between weather parameters and the number of new cases. As sketched on the figure above, we want to provide 3 graphs, a first one (**Fig.1**) that displays the cross correlations between weather parameters and the number of cases for different time lags (indeed the effect of wind for example may have an impact on the number of new infections only few days later)(*for calculating the correlation we will rely on an existing lib https://github.com/agrueneberg/Spearson*). Then, if **Fig.1** shows any interesting correlation you can choose the desired weather feature with a button and plot the corresponding signals on **Fig.2**. This plot has two time lines, one for the number of cases and the second one for the selected parameters, indeed : it should let the ability to slide the weather signal from the number of cases to visually observe any correlation between the signals (*Lecture and exercise session on interactive d3.j applied to time-series for interactions, D3Brush,D3Zoom*). Finally, **Fig.3** shows the correlations for the selected feature by province (*Lecture on maps*). Some additional widgets are also foreseen: **Fig.2** should let the ability on the basis timeline to choose a given time period and update **Fig.1** and **Fig.3** for this period. **Fig.3** should also let the ability to select a province, so that we can observe more specifically correlations on **Fig.1** and **Fig.2** for the selected province.