December 10th, 2020

Baltzi Sofia

Aueb

Pandemic illustrated 2.0

Assignment 2 for Data Visualization and Communication

# **Introduction**

A novel coronavirus, later named Covid-19, was detected in Wuhan, China in 2019. By November of 2020, 218 Countries and Territories around the world had reported to have confirmed cases of the virus.

For the first assignment, we used data\* from the Open Data Portal of European Union, in order to visualize the worldwide spread of the virus. For this assignment, we will try to enhance the value of some of the static visualizations, by adding interactive features. Figures 1.2, 2.2, 3.2, 4.2 were created using Tableau. Figure 5 was created using R and Shiny library.

\*Note that data was available for up to November 21st

# **Worldwide Spread**

Figure 1.1 is a static representation of Coronavirus spread over time per continent, as percentage of continents’ population. Each continent is represented with a different color.

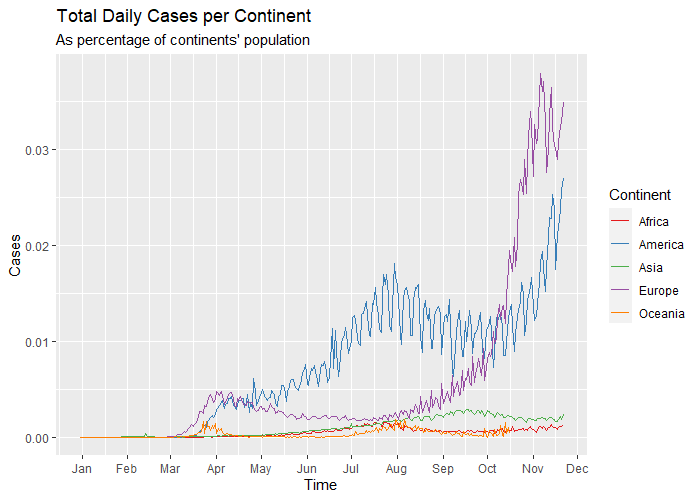


Figure 1.1: Continent spread - Static

Figure 1.2a is an attempt to make this exact same diagram interactive, in order to add some value to the representation. Firstly, when hovering above a line of the **line plot**, a **pane** comes up that shows which continent is selected, for which date and what the percentage cases/population was on that date for the specific continent. In the static plot, one did not have that information and could only speculate by observing Time and Cases axis.

Additionally, when clicking on a point of the line plot, a **treemap**\* of the countries of the continent for the specific date comes up, that shows the breakdown of cases within the continent. When hovering over a country in the treemap (Figure 1.2b), a **pane** like the one in the line plot comes up as well. The treemap provides with extra information that helps to explain the numbers of the continents.

If we wanted to implement the dashboard of figure 1.2 in a static form, we would need one line plot and 5 treemaps, meaning that the interactive dashboard saves space, time and makes for a cleaner look since it is not filled with labels.

\*Color scale was not mandatory, but it made the differences between boxes clearer.



Figure 1.2b: Continent spread - Interactive

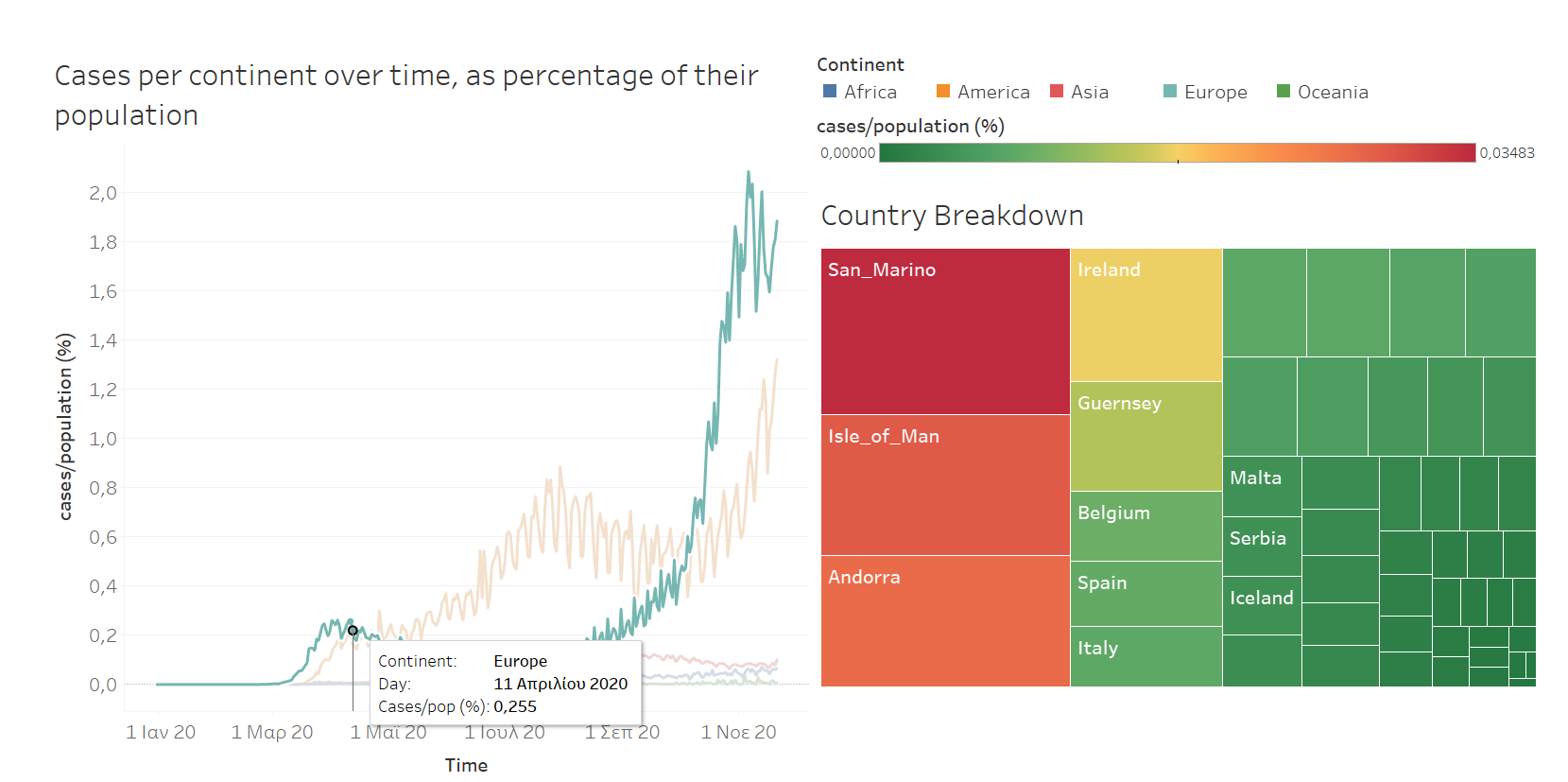


Figure 1.2a: Continent spread - Interactive

# **Europe**

In Figure 2.1, we see a map of Europe, were countries are colored according to the percentage of their population that got infected with Covid-19 the last 30 days of data. Countries with small infection percentage are colored blue, and those with large are colored red.

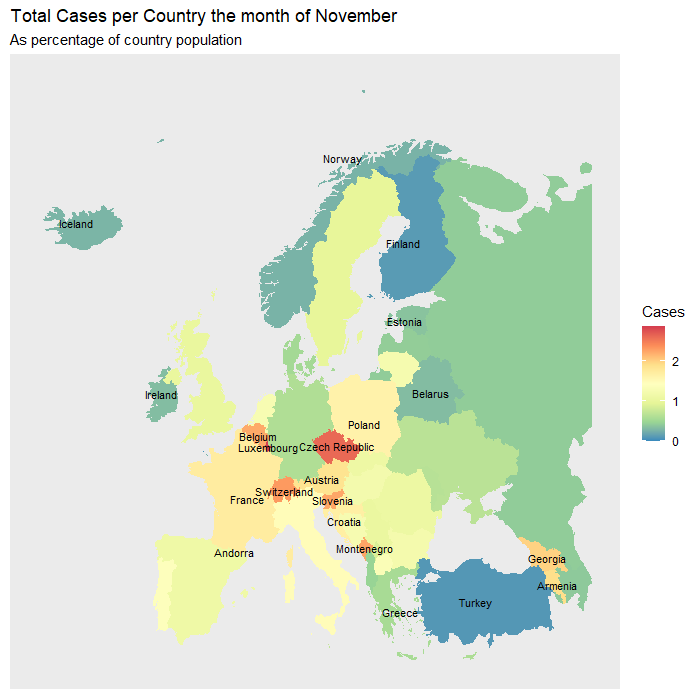


Figure 2.1: Europe spread - Static

Figure 2.2a is an interactive alternative for Figure 2.1. The difference is that the color scale is replaced with different sized bubbles that represent the magnitude of the metric plotted (cases/ population percentage). The left **map** depicts the cases and the right one the deaths.

Instead of plotting just the data for the last month, the data is grouped by month, and the user can choose the month to plot utilizing the **slider/ drop down list**. Therefore, we do not need one plot for each month we want to see, meaning that the interactive representation is more inclusive and informative. The **tables** show the values of the bubbles.

Figure 2.2b shows what happens when the user clicks on a country. In both cases’ and deaths’ maps the chosen country is highlighted and the tables bellow are reduced to showing only the target’s metrics. This utility helps the user to focus in one country and know the exact metrics rather than an abstract approximation that comes from the color coding of the static plot.

Figure 2.2c shows that the user can choose more than one country in order to compare the metrics between these countries. Again, that comparison in the static plot could only be performed in an abstract way by comparing color hues and only for the plotted month.

Figure 2.2 b: Europe spread - Interactive

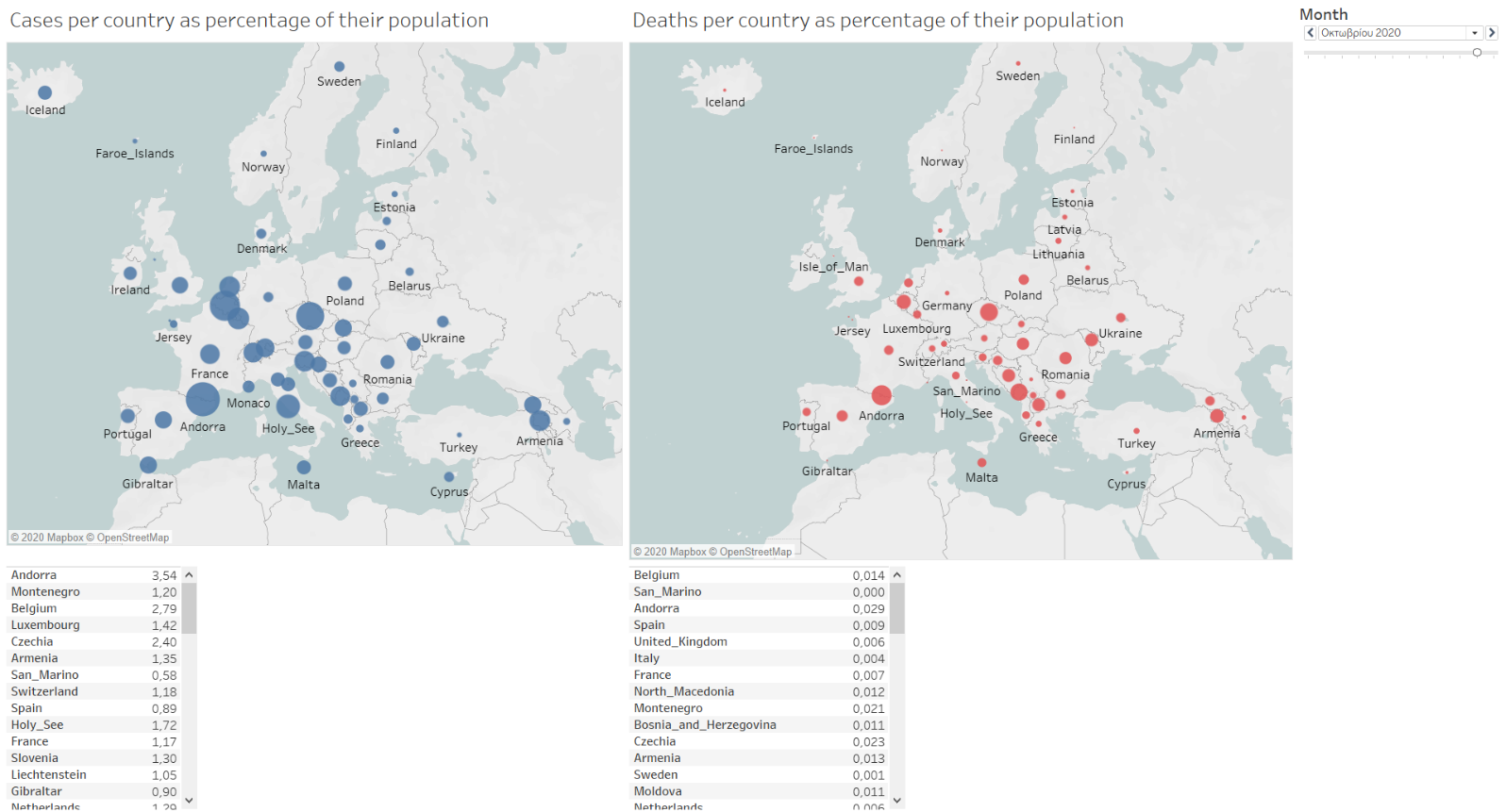


Figure 3.2 a: Europe spread - Interactive

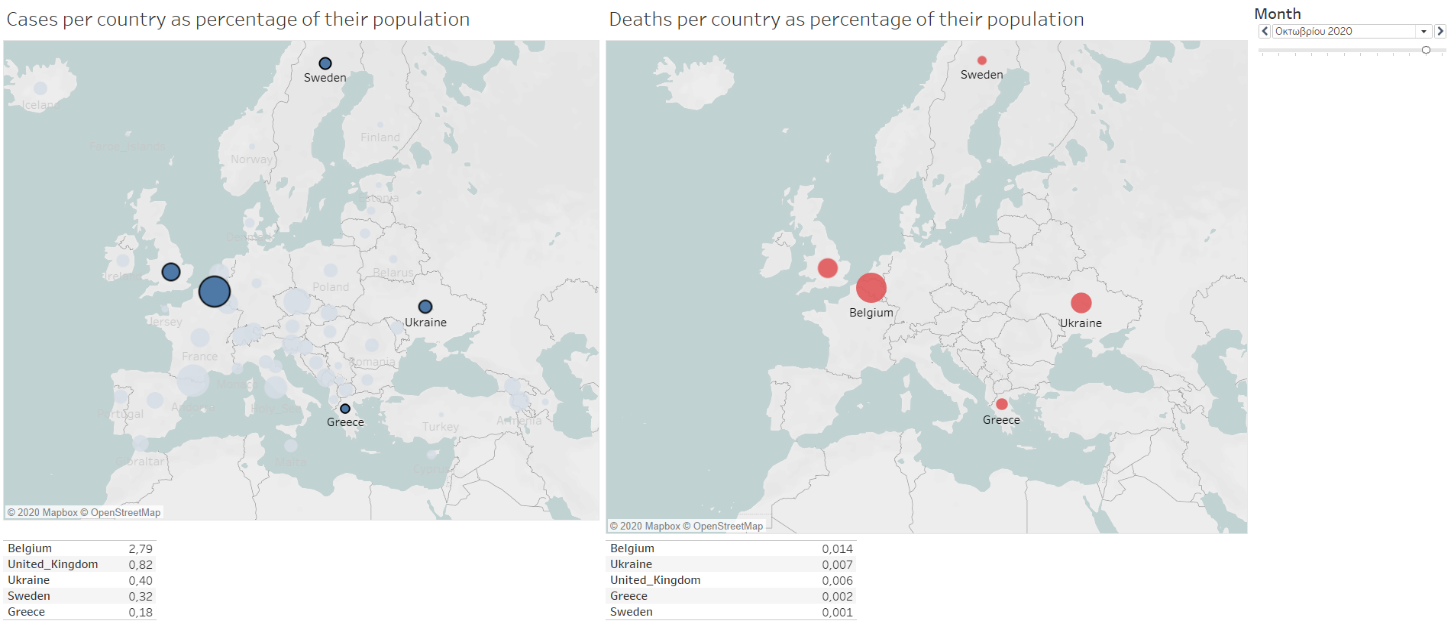


Figure 4.2 c: Europe spread - Interactive

# **Greece**

Figure 3.1 highlights the ranking of Greece among the other countries of European Union in terms of cases.

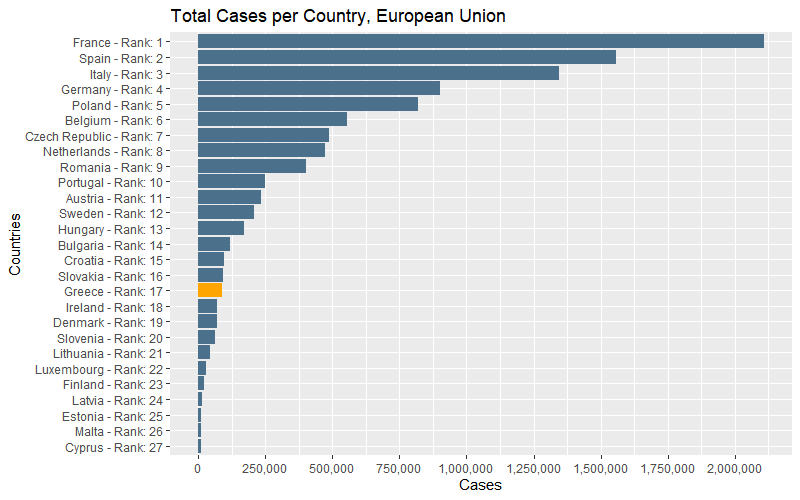


Figure 3.1: Ranking in terms of cases for European Union - Statics

That ranking is shown on the **treemap** of the interactive Figure3.2 (the difference is that we now have as a metric the cases/population percentage). By hovering over the treemap, we can see a **pane** with the exact value of the metric for each country, which we could only approximate in the static plot.

Another added feature in this representation is that the treemap acts as a filter in order to compare Greece’s monthly metrics with other countries’ metrics. On clicking a country, the **top right bar plot** is updated to show the monthly values of the target country, whereas the **bottom right plot** always shows the values for Greece. Therefore, the same dashboard can be also used for a comparison between countries in a monthly level.

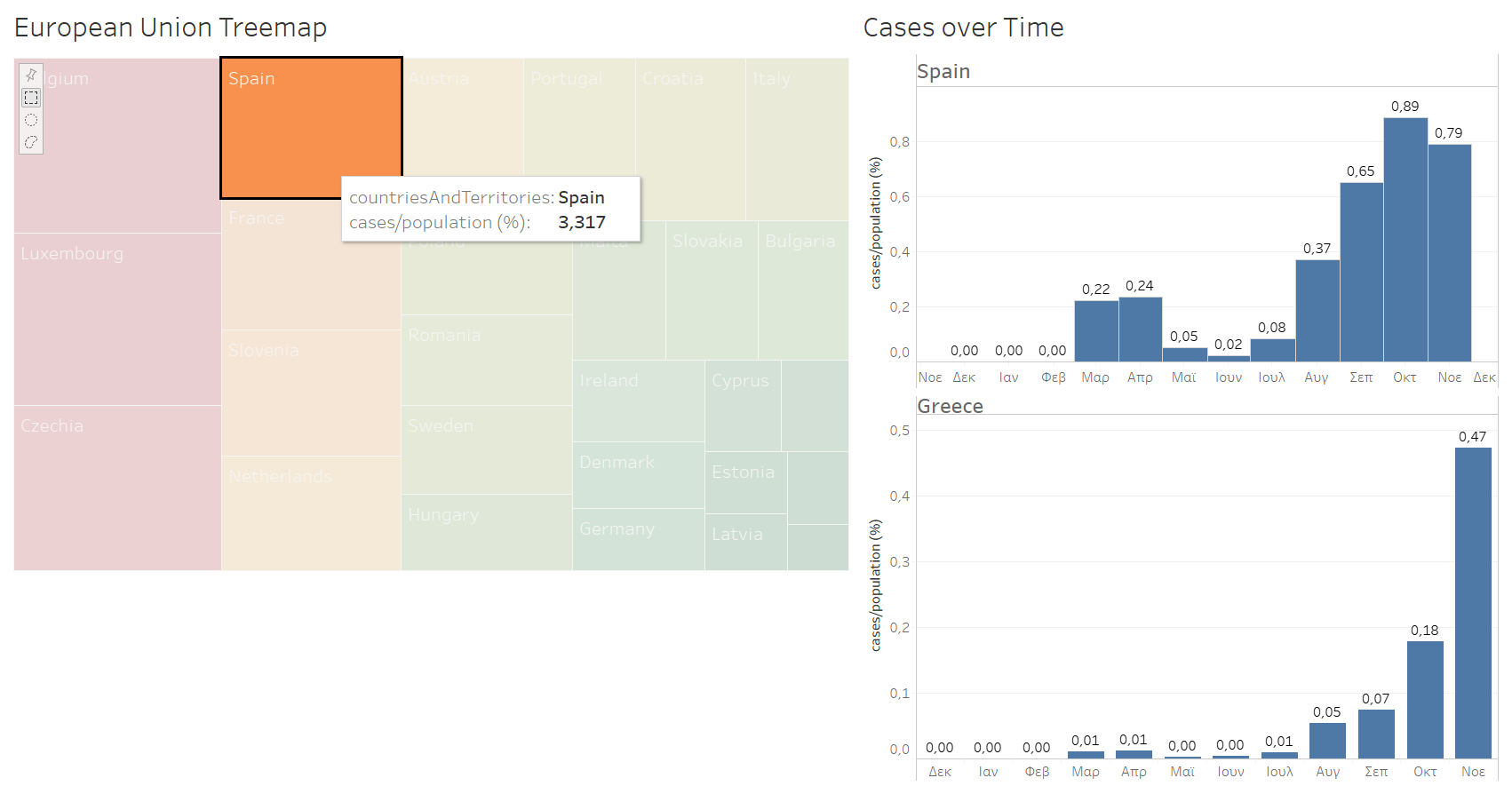


Figure 3.2: Greece/EU comparison - Interactive

Figure 4.1 was used to present monthly cases/ deaths for Greece.

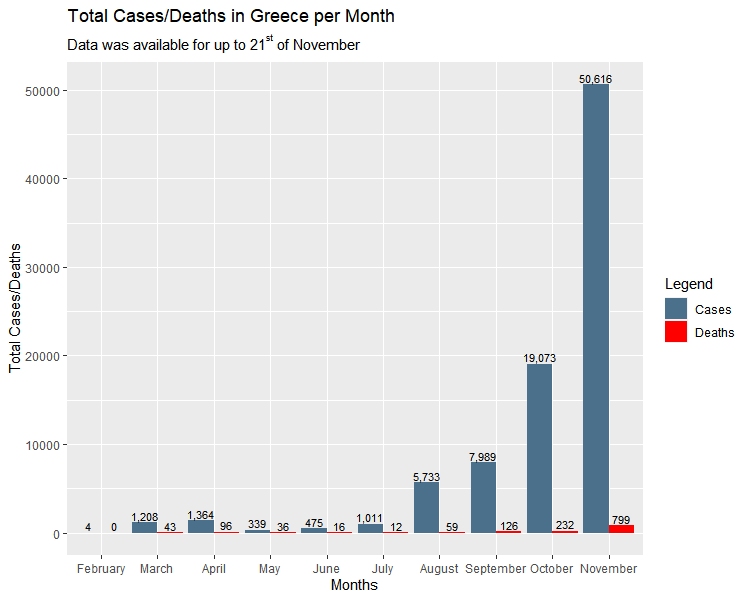


Figure 4.1: Greece’s Breakdown - Static

Figure 4.2a is the interactive version of the same Figure. The bars of cases and deaths are presented in **two separate diagrams** for a clearer view. In the general form the Day Drill Down part of the Figure consists of **two line** plots, one for cases, the other for deaths, that include all dates of data. When hovering over a point, a pane comes up that shows the number of cases and deaths.

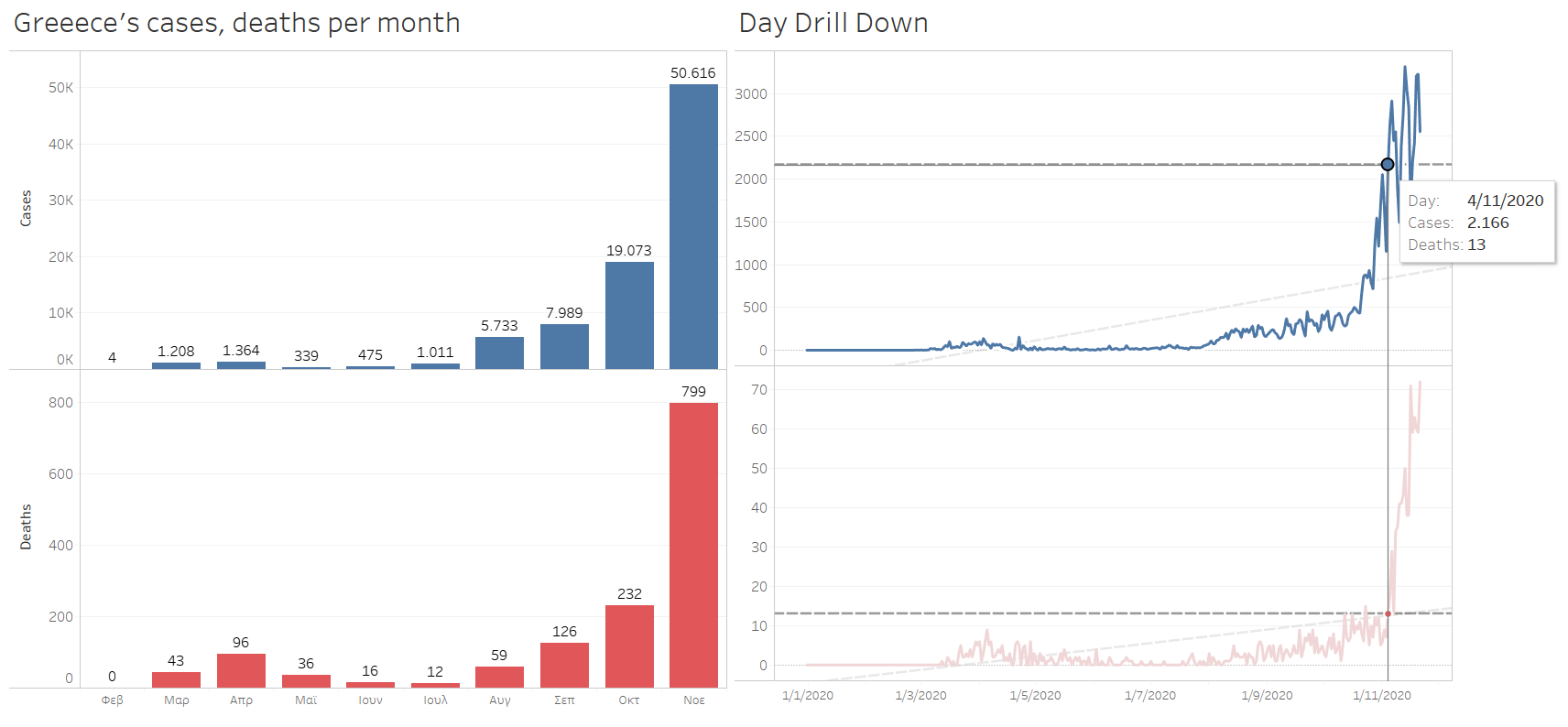


Figure 4.2a: Greece’s Breakdown - Interactive

An added feature is that the user can **choose a month** in the bar plots and see the daily breakdown in the line plots, as it can be seen in Figure 4.2b. That is useful for when the user wants to focus on specific dates. In the static plot one could only compare the months, one with another. We would need 11 static plots for the 11 months of data to get the same value as with the interactive one. The gray dashed lines are trend lines for daily cases/ deaths, and they are also updated when the filtering is updated. By comparing the trend lines, we can see if the trend of deaths follows the trend of cases.

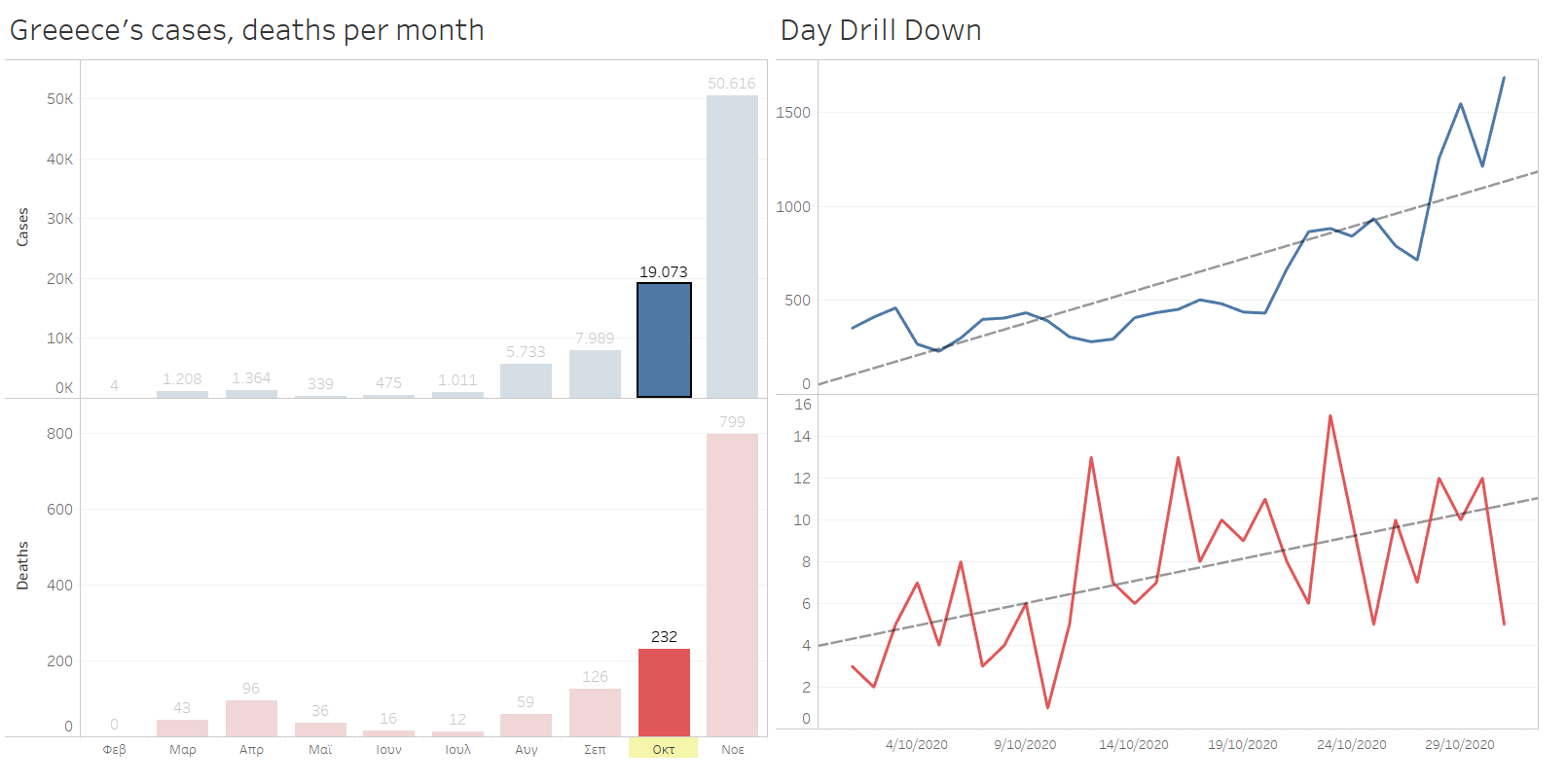


Figure 4.2a: Greece’s Breakdown - Interactive

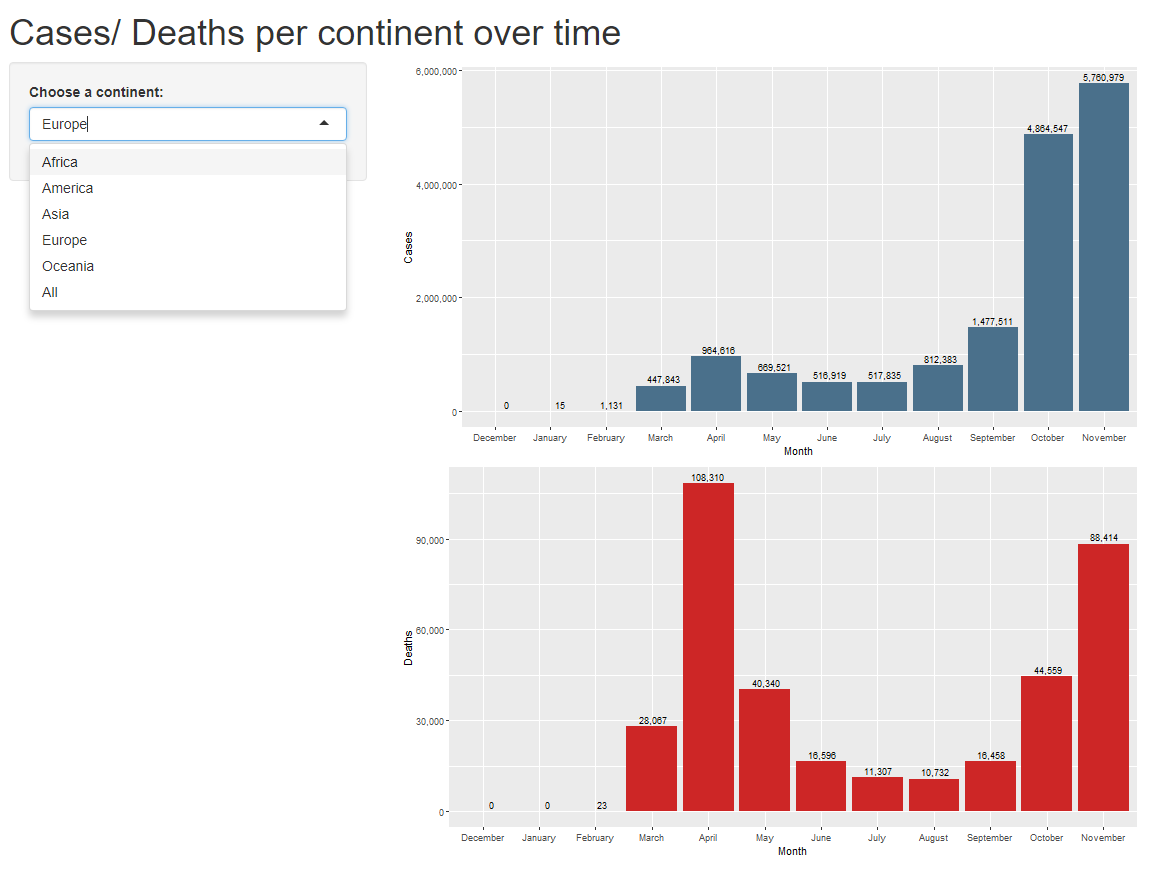
**Worldwide Spread (Shiny)**

Lastly, Figure 5 shows an interactive diagram made with R and library Shiny. The drop-down list on the left can be used to filter a continent. For this continent two bar plots are created, one showing the total monthly cases and one the total monthly deaths.

The static version of this plot would consist of six plots for cases (five for the continents and one for the total) and six plots for deaths (five for the continents and one for the total). Another static version would consist of two plots having six bars for each month (five for the continents and one for the total). The first version would take too much space, while the second would make for a busy plot.

Therefore, the interactive plot gives a clean look that does not take too much space.

Figure 5: Cases/ Deaths per continent - Interactive



**Conclusion**

We can use our observations about static and interactive visualizations in order to extract some benefits of the latter:

* They can make our plots less busy by removing labels that can be shown with hover over.
* They can be used to save space in our dashboards (e.g. N static plots versus one interactive with N options to show).
* They can be used to incorporate more information.
* They can be used to tell a story with the data.
* They can be used for comparisons, to spot relationships and trends.
* They can be customized by the designer as well as the user.

Dynamic visualizations do have quite a lot of benefits but should be used when they truly have to add something to our analysis, otherwise they will just overcomplicate the designing process and will confuse the reader.