Covid 19 - Analysis of the Italian data at the regional level

Lorenzo Crotti, Ilaria Prometti, Paola Rossi, Riccardo Valle

Last updated: April 2, 2020

We will consider the following variables:

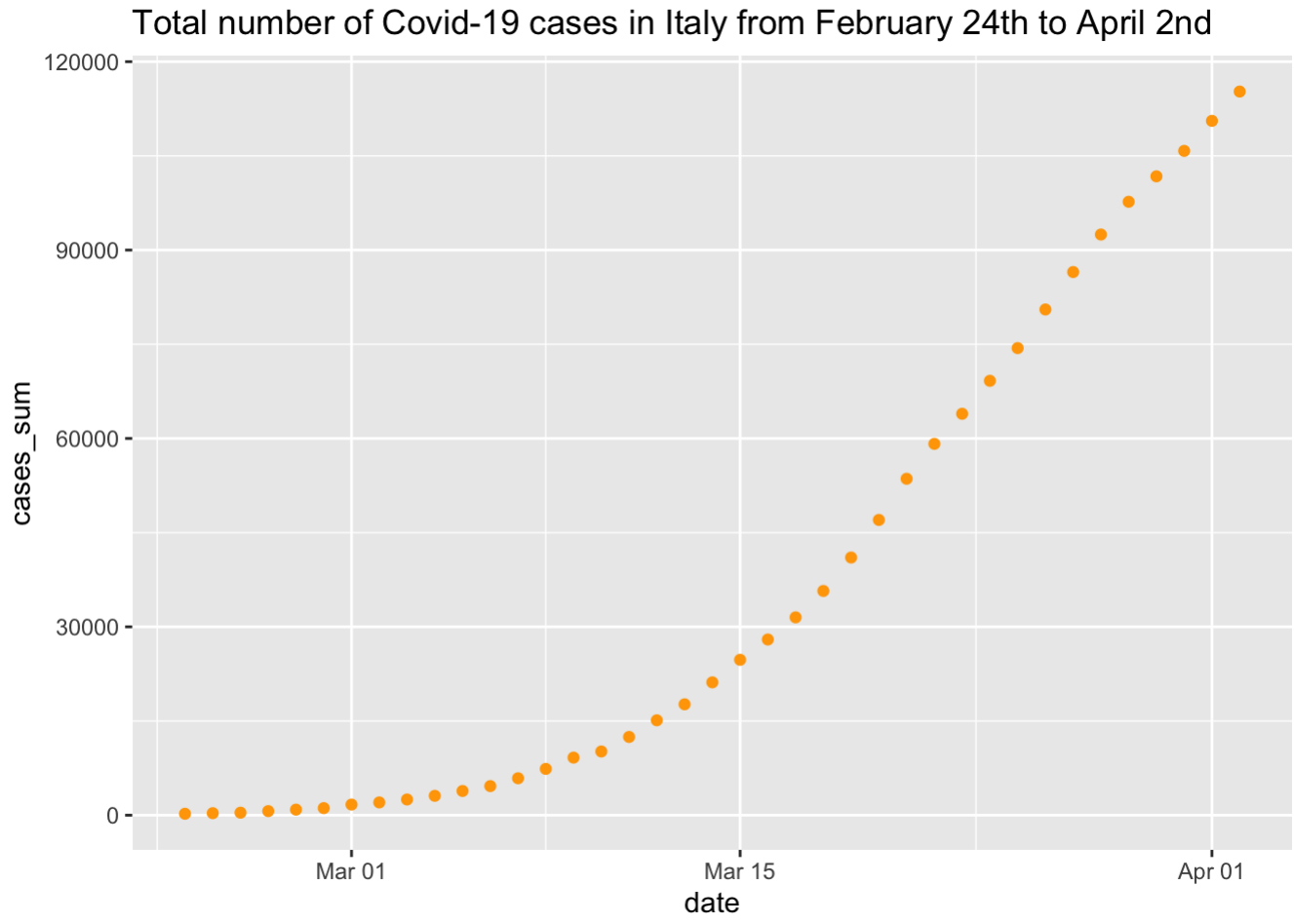
- number of Covid-19 cases,
- number of hospitalized persons,
- the number of healed persons,
- the number of deaths,
- the number of swaps

at the national level and separately for each region during the time period from February 24th to April 2nd.

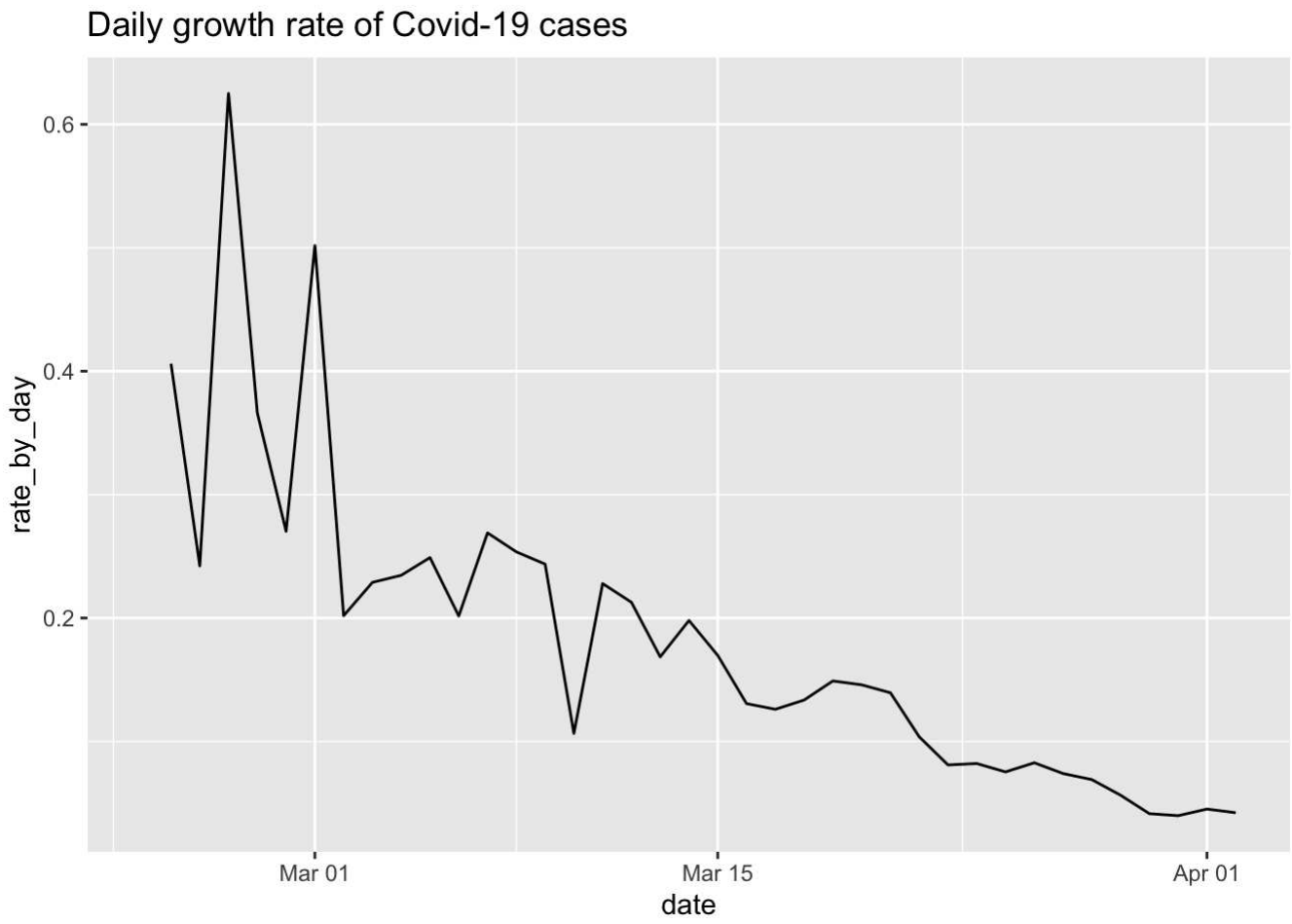
National level

Let's start by reporting some statistics at the Italian level.

The following plot displays the total number of cases for each day from February 24th to April 2nd. We observe that on April 2nd in Italy the total number of cases is 115242.



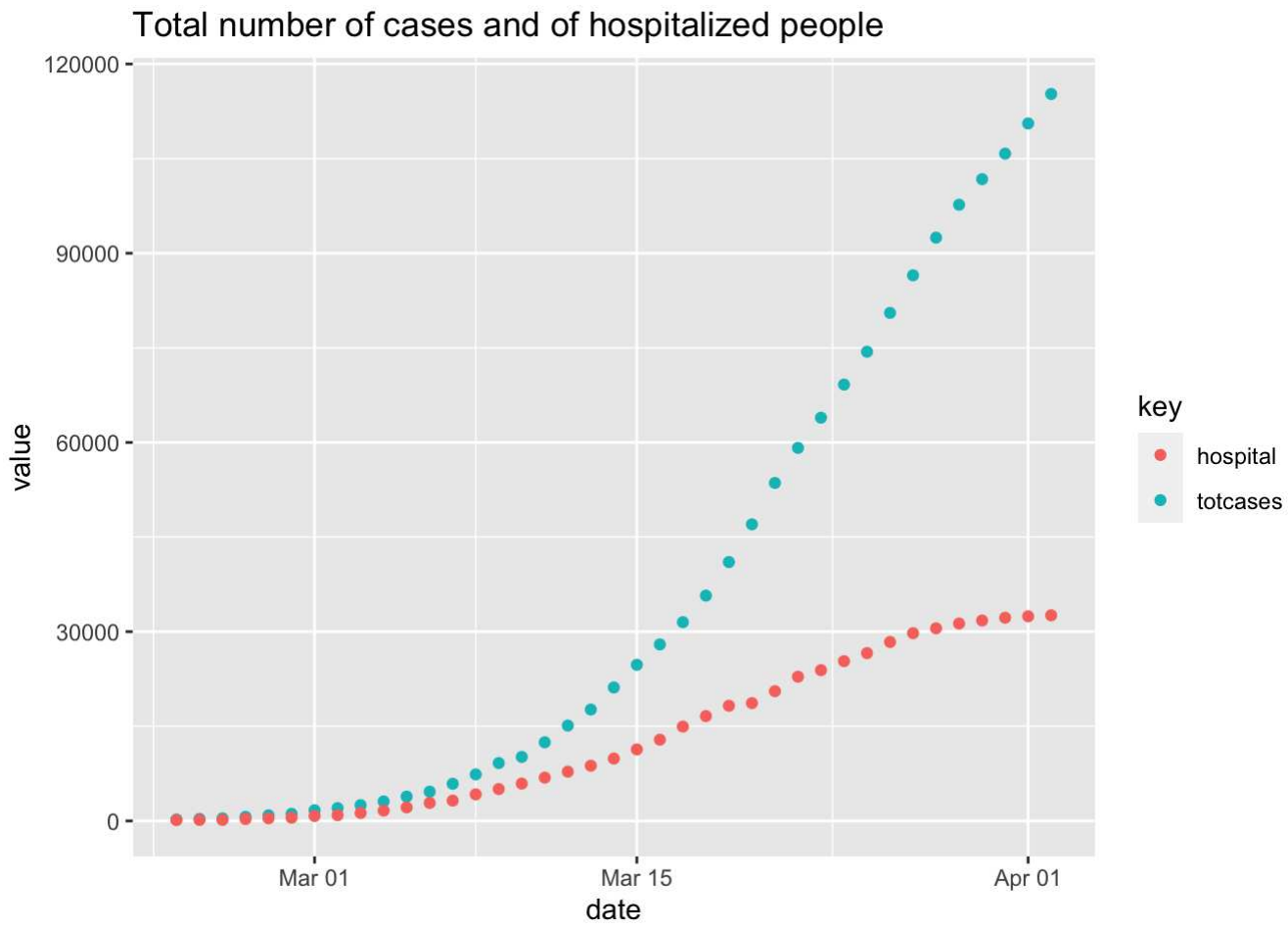
Over the considered time period the number of Covid-19 cases has increased rapidly. However it seems that this growth has been slowing down in recent days. This is visible from the following chart representing the daily growth rate in the number of cases:



The total number of cases is given by the sum of:

- number of positive people in hospitals,
- number of positive people at home,
- number of survived people (healed),
- number of deaths.

The following plot represent the total number of cases and of hospitalized people. The graph shows that as the number of contagion increases not everybody has been hospitalized.



Regional level

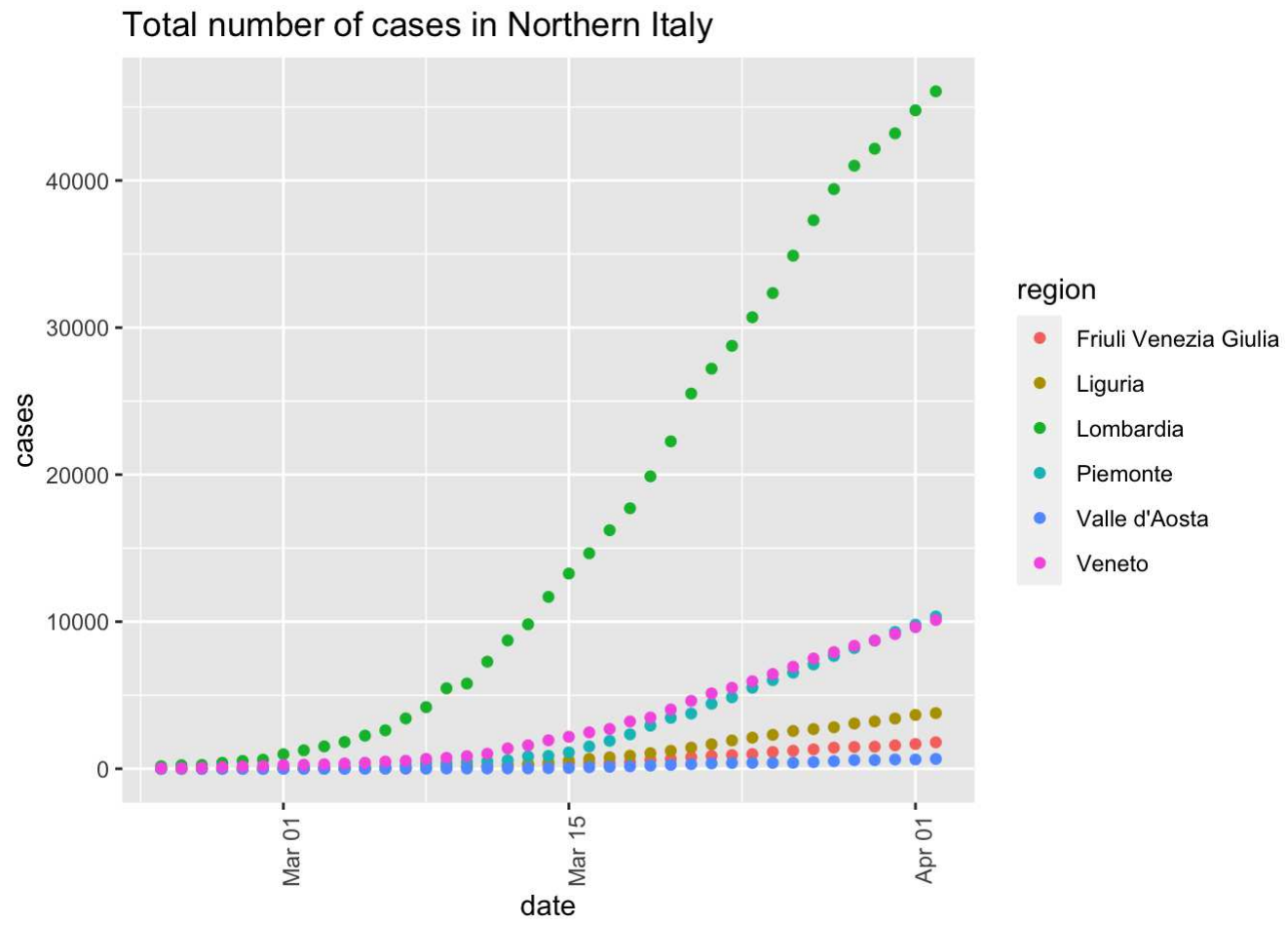
The following table reports the distribution of the total number of cases by regions (there are 21 spatial units because “Trento” and “Bolzano” are considered separately):

```
## # A tibble: 21 x 2
## # Groups:   region [21]
##    region                cases
##    <chr>                 <dbl>
##  1 Lombardia             46065
##  2 Emilia-Romagna        15333
##  3 Piemonte              10353
##  4 Veneto                10111
##  5 Toscana                5273
##  6 Marche                 4098
##  7 Liguria                3782
##  8 Lazio                  3433
##  9 Campania               2456
## 10 Puglia                 2077
## 11 P.A. Trento            2003
## 12 Friuli Venezia Giulia  1799
## 13 Sicilia                1791
## 14 Abruzzo                1497
## 15 P.A. Bolzano           1479
## 16 Umbria                 1128
## 17 Sardegna                794
## 18 Calabria                691
## 19 Valle d'Aosta           668
## 20 Basilicata              246
## 21 Molise                  165
```

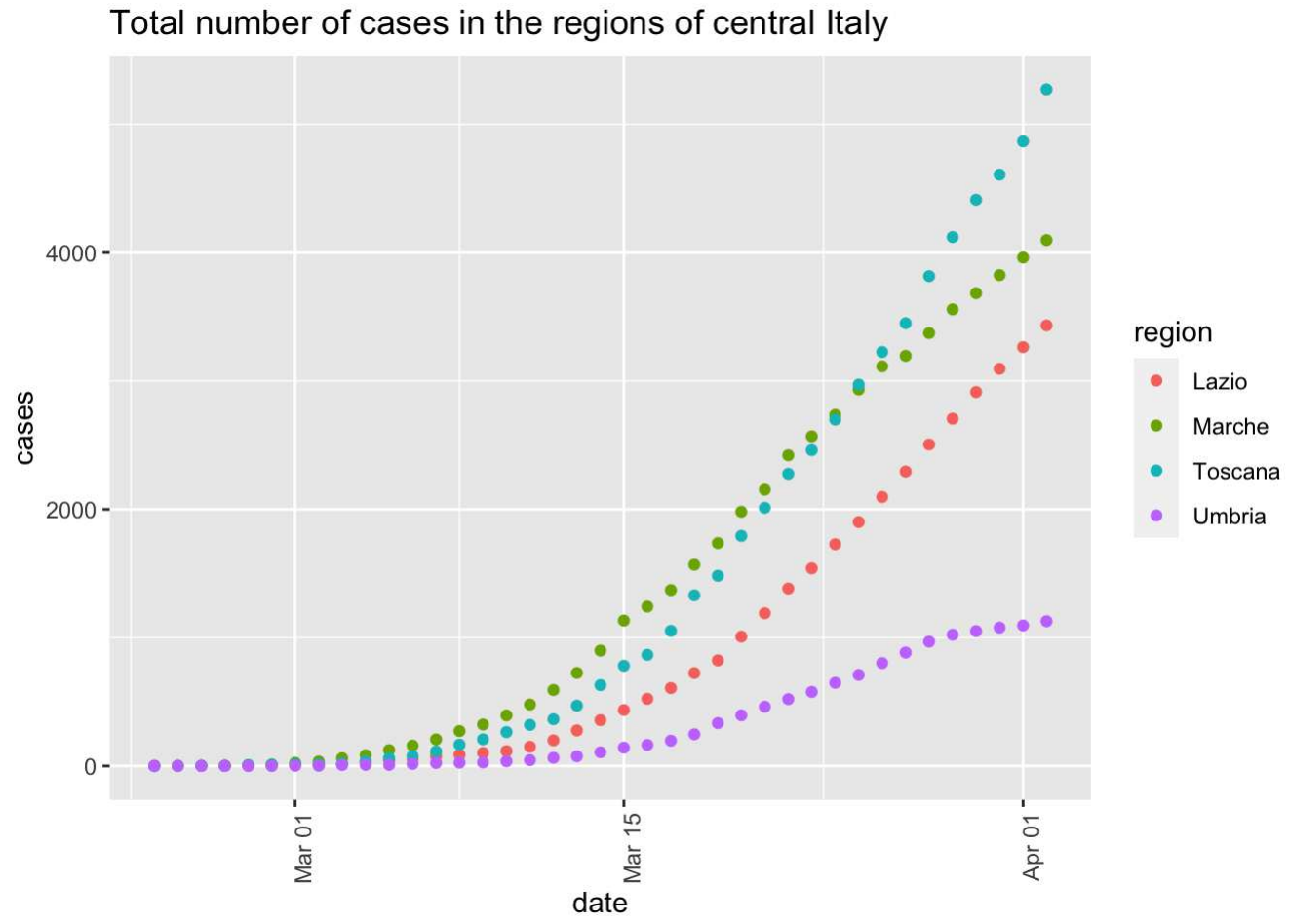
On April 2nd the region Lombardia records the highest number of cases followed by Emilia Romagna and Piemonte.

For a clearer graphical representation we divide the Italian regions 3 macro regions (North, Centre and South) and we represent the total number of cases.

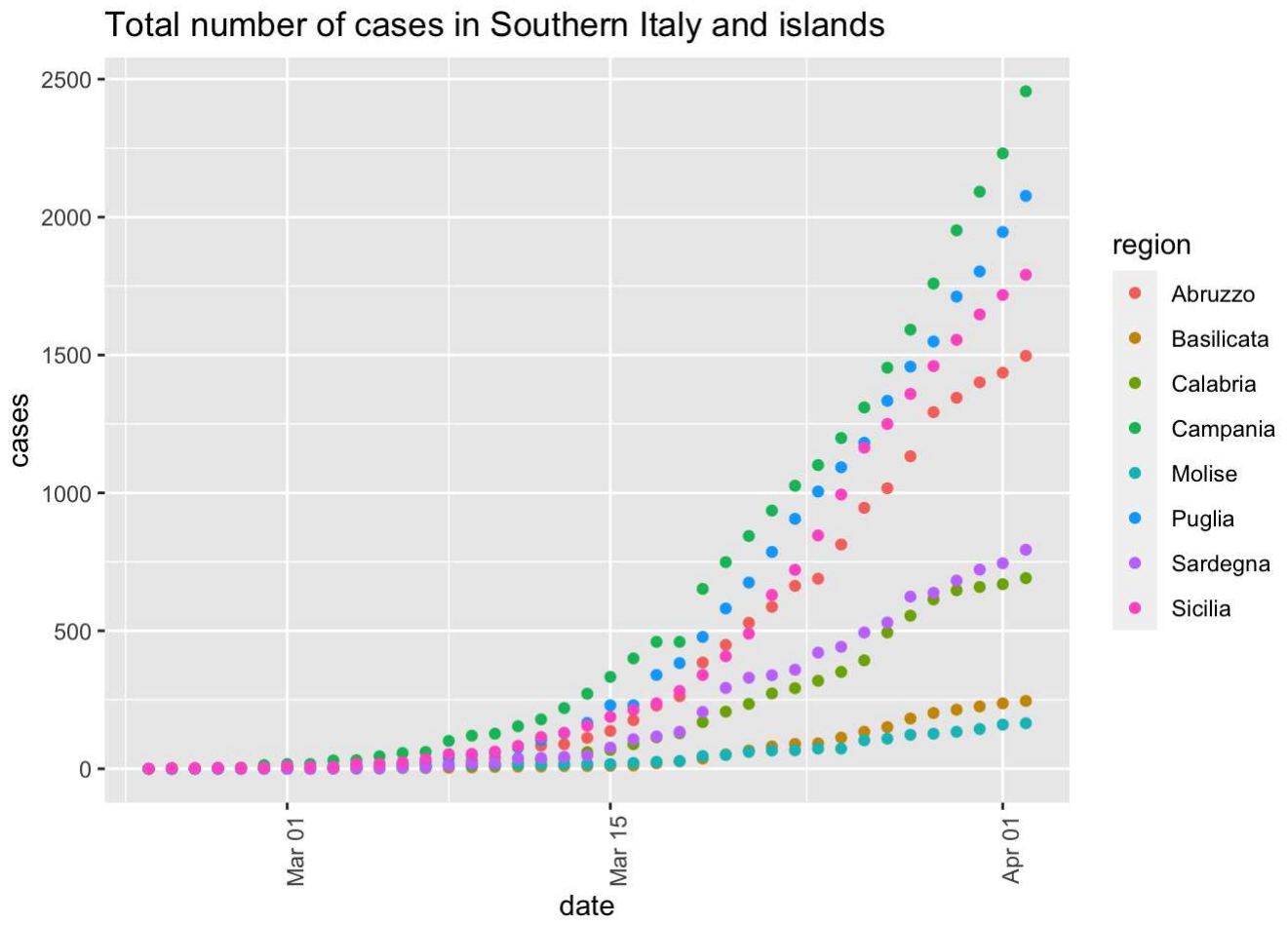
NORTH



CENTRE



SOUTH

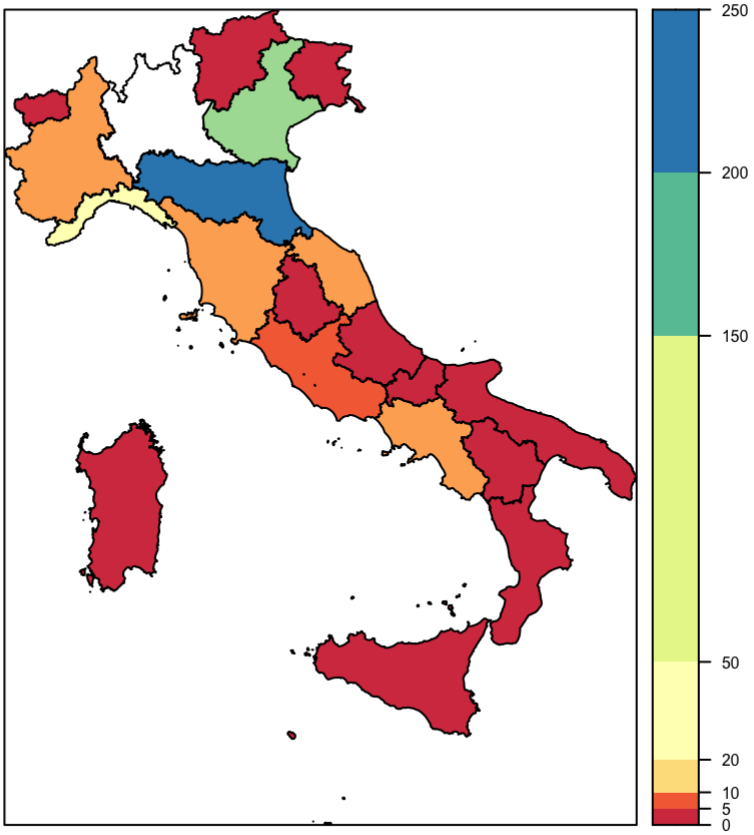


We want to represent the growth of the total number of cases during the considered time interval To do this we consider three dates: February 29th, March 14th and April 2nd (the last available date). For the analysis we sum the data related to Trento and to Bolzano provinces and we obtain the value for the whole Trentino-Alto Adige region.

First date: **February 29th**

In the following plot we decided to exclude the region Lombardia because it has a much higher number of cases than in other regions. The number of cases in Lombardia is shown below the graph.

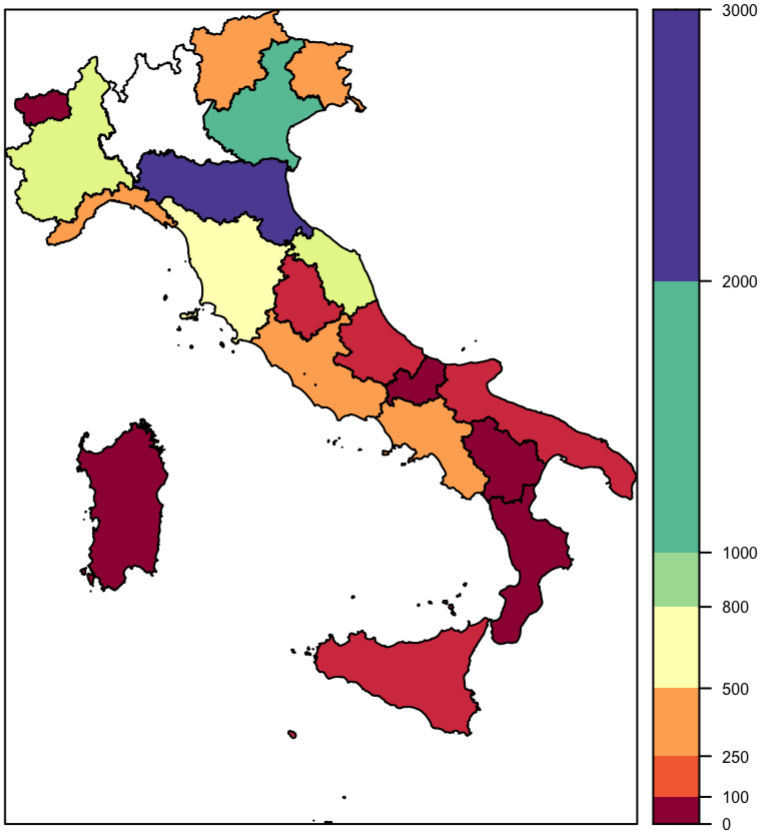
Distribution of Covid-19 cases by region, as of 29 February 2020



Cases in Lombardia: 615

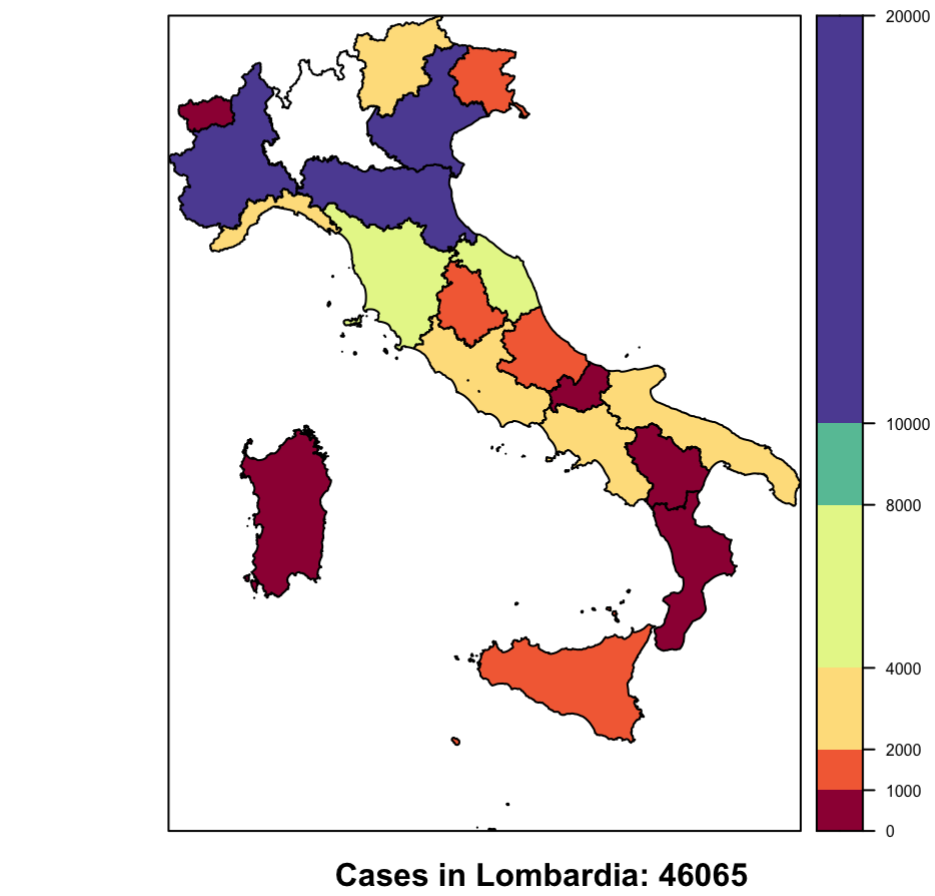
Second date: **March 14th**

Distribution of Covid-19 cases by region, as of 14 March 2020



Cases in Lombardia: 11685

Third date: **April 2nd**



SWAPS AND MORTALITY RATE

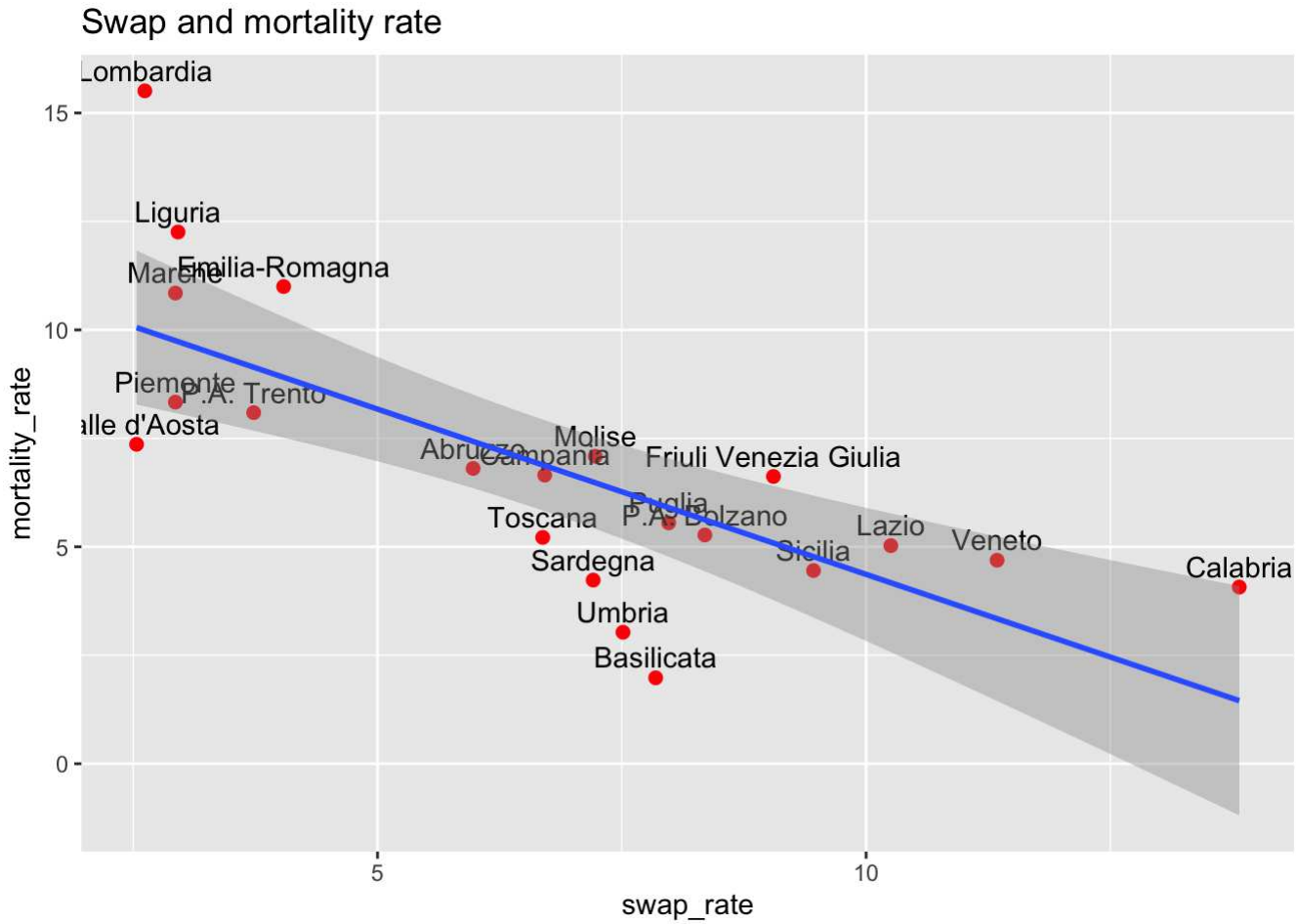
We want to show that for some regions a relative lower number of swaps has been used to test Covid19 with respect to others. We think that in these regions the mortality rate will be on average higher because these regions use swaps for patients with strong “symptons” while in other regions, where an higher number of swaps has been used, in the variable “cases” are considered even patients with weaker symptons. For this reason we expect a lower level of the mortality rate.

Let’s study the correlation between the swap rate (no. of swaps / no. of cases) and the mortality rate (no. of deaths / no. of cases).

```
##
##  Pearson's product-moment correlation
##
## data:  covidnew$swaps_to_cases and covidnew$mortality_rate
## t = -4.6655, df = 19, p-value = 0.0001686
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.8836561 -0.4368035
## sample estimates:
##        cor
## -0.7307066
```

As we expect there is a negative statistically significant correlation.

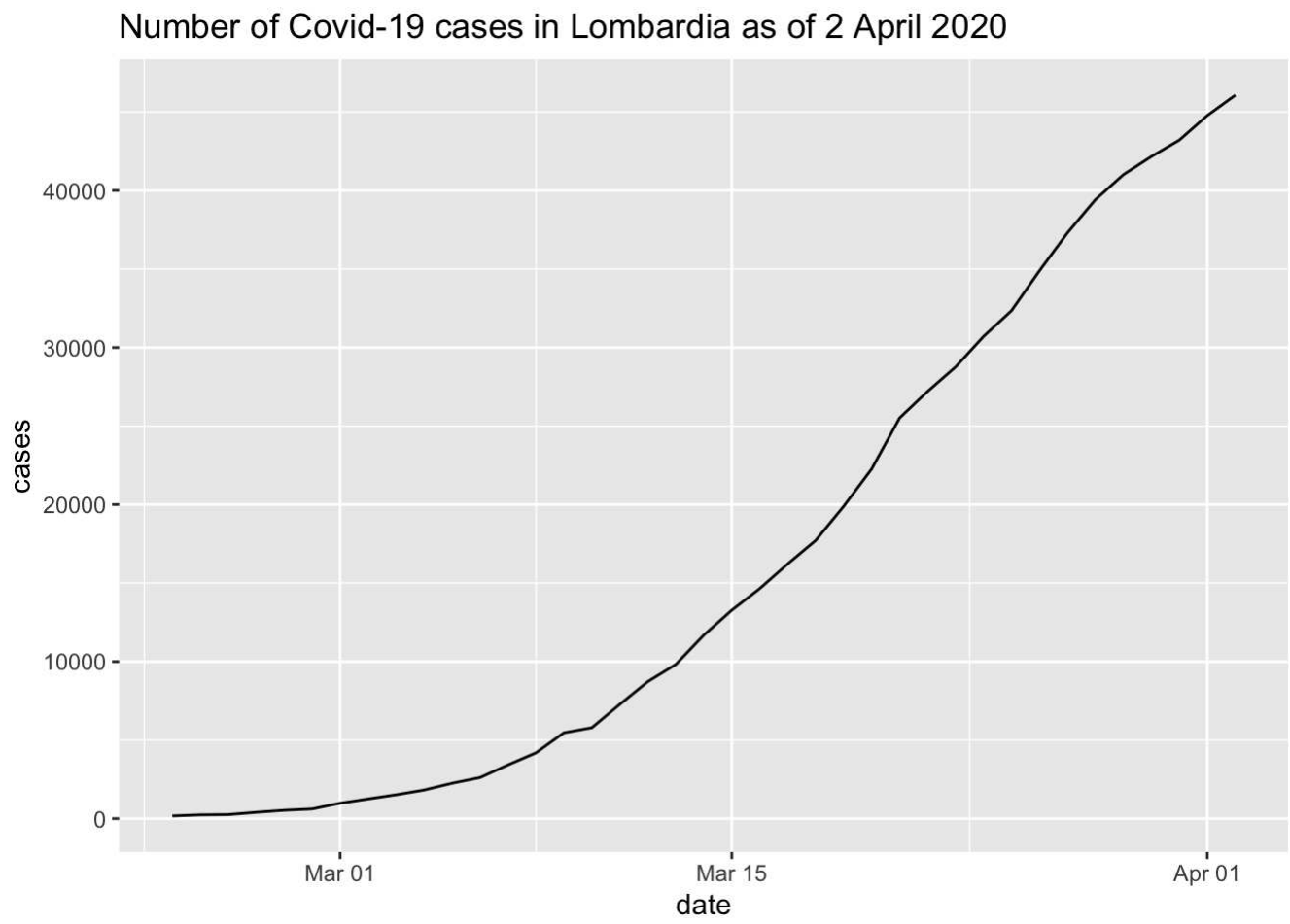
The following graph shows this negative relationship:



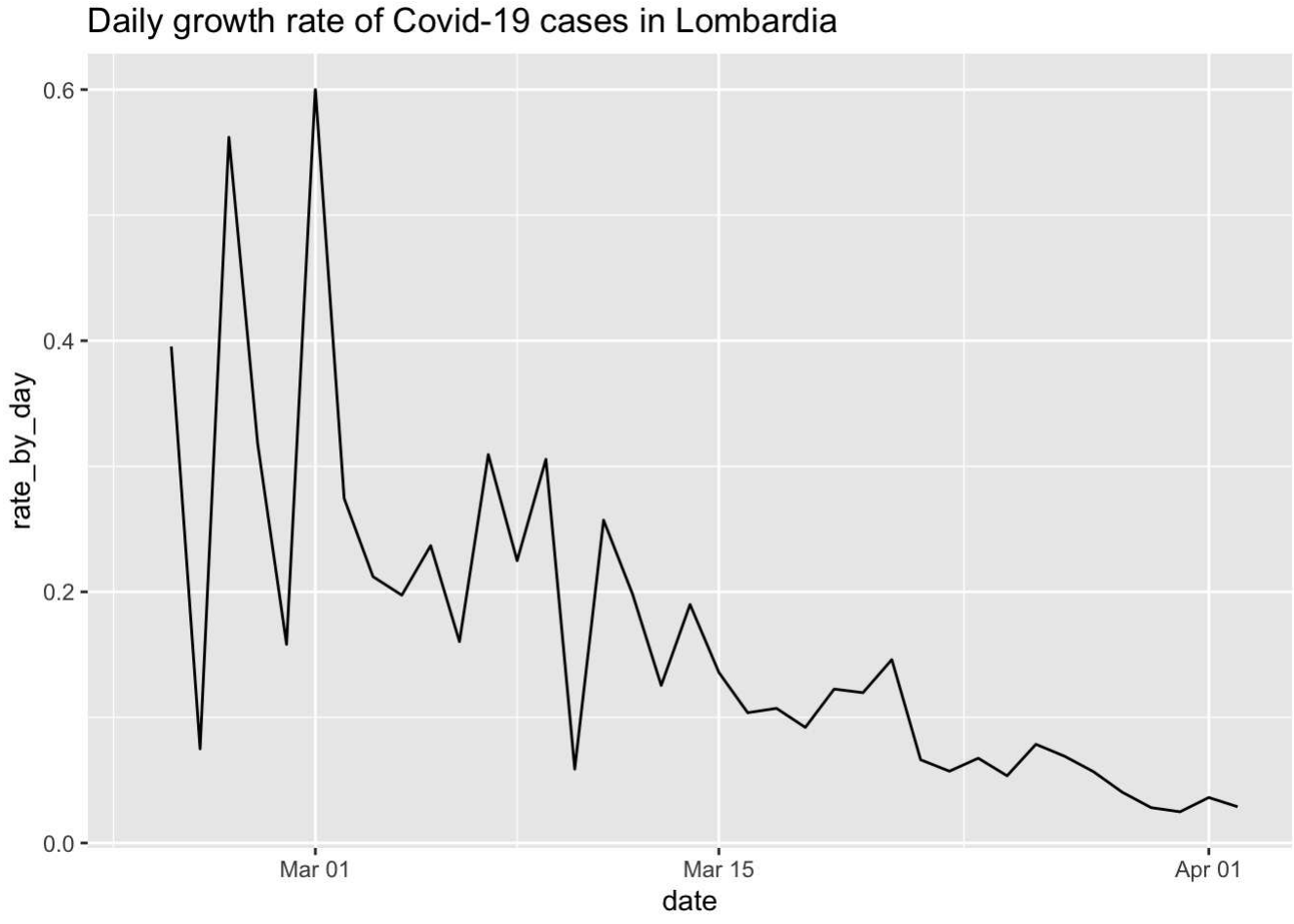
We can see that as the swap increases, the mortality rate decreases. The region Lombardia which has carried out fewer swaps than for example the Veneto region shows an higher mortality rate.

Lombardia

We wonder if contagions are increasing or decreasing in Lombardia. The following plot displays the total number of cases for Lombardia region. The region has experienced an exponential growth in the number of cases, but in recent days this growth seems to have slowed down.



In this regard the following plot shows the daily growth rate in the number of cases in Lombardia.



On April 2 in the region Lombardia about 1300 more cases were registered than on the previous day. A peak can be noted on March 21 when the number of new cases compared to March 20 is around 3200. Instead, in the following days the growth was lower. There is another peak between 25 and 26 March (+ 2500 new cases). However in the last few days the growth has started to decline again.