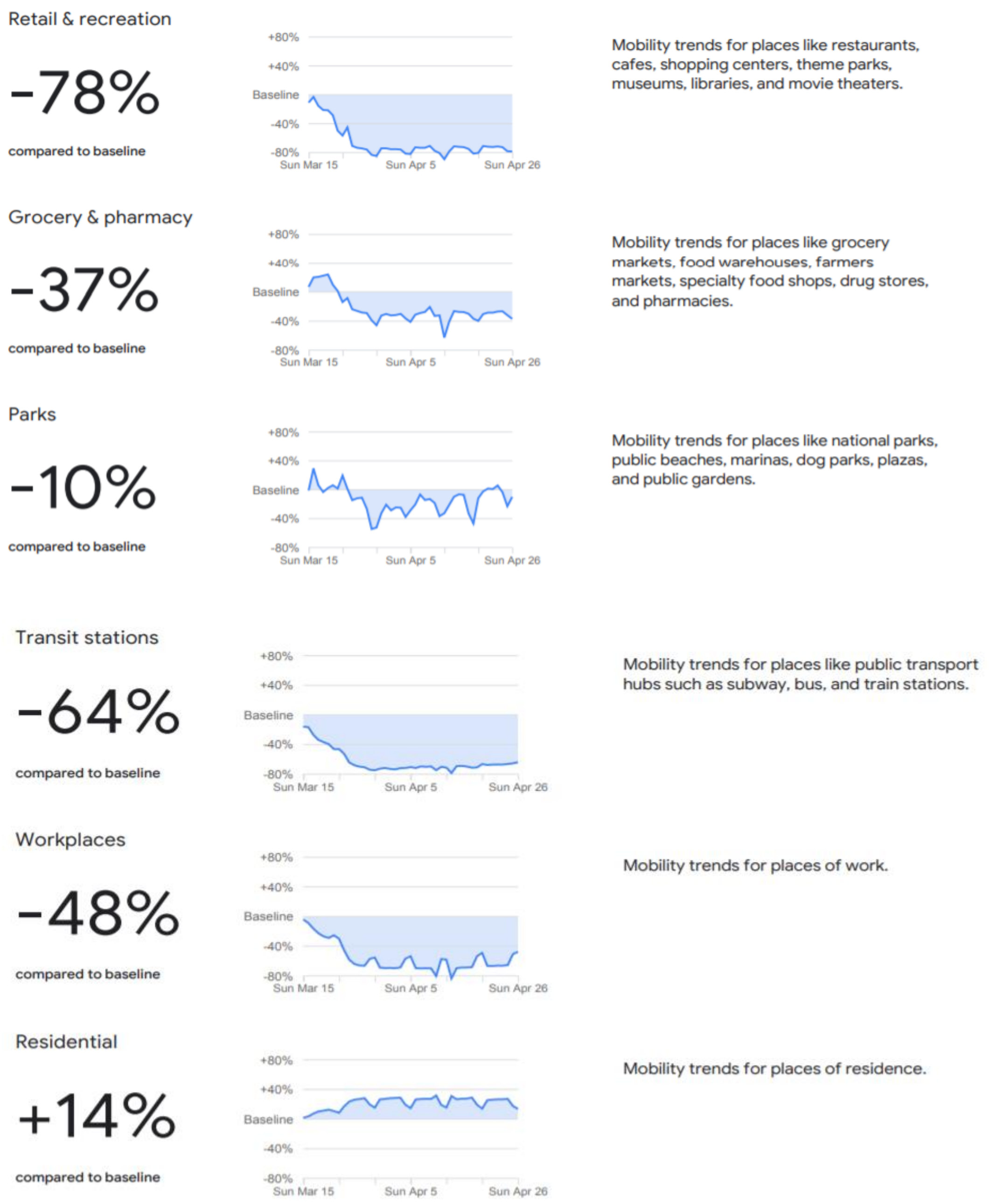
COVID-19 reproduction number and Google mobility data

The main goal of this exercise is to find relation between changes in COVID-19 reproduction rate and changes in average society movements measured by Google through data collected by Google Maps.

Reproduction rate indicates how many healthy people on average will an infectious person infect during infectious period.

Google mobility data is divided in 6 categories, as shown in pictures below (data for UK).



Source: https://www.google.com/covid19/mobility/

Categories Residential and Parks have been excluded from calculation as those places are not considered to be significant for disease spreading (because people staying at home or going to park instead of going to work/bars/gyms actually helps in lowering the reproduction number R).

Average mobility was calculated as an average of the remaining four categories.

Two different approaches were taken when calculating **reproduction rate R**:

1. R calculated using the official numbers of Confirmed cases, Recovered cases and Deaths. It is only a rough approximation for R. The main assumption is that new infections caused by actively infected people today are visible in official numbers during next 8 days (because infectious period is 3 days on average and there is also an incubation period of 5 days).
   Basically, formula for this would be:
   R (t) = (Confirmed Cases (t + 8) – Confirmed Cases (t)) / (Active cases (t))

2. R calculations made by The LSHTM (London School of Hygiene & Tropical Medicine) using more complex methods and additional information and data. Time series for R can be found on the following link: https://github.com/epiforecasts/covid-global/blob/65f4b0b38526e66d9b59fea7d6413f3929b290f3/national-summary/rt.csv

R movements for UK and Croatia calculated by both approaches are shown in Figure 1 and Figure 2.

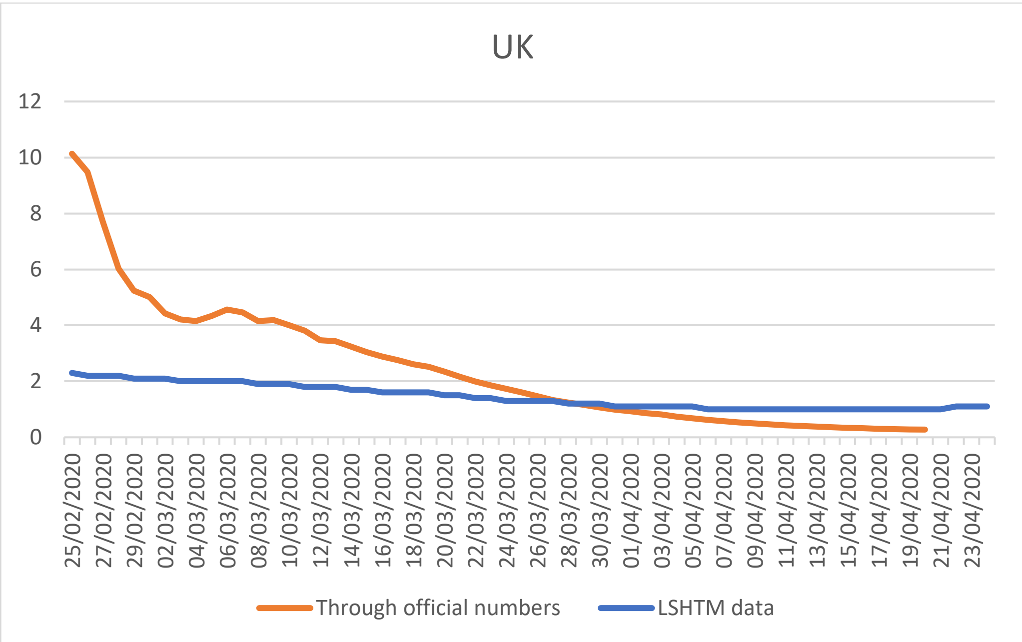


Figure 1

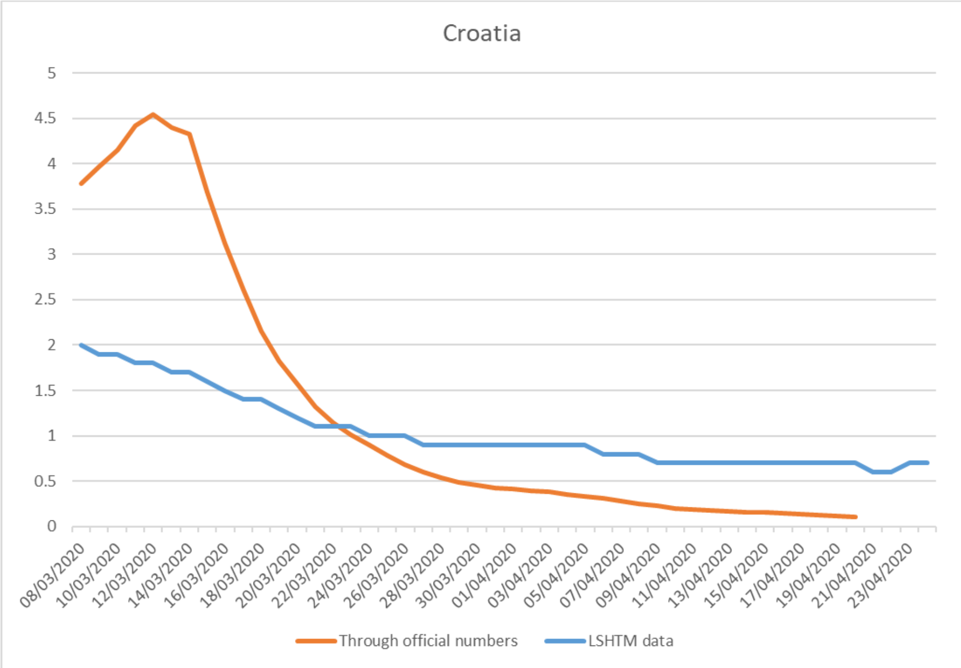


Figure 2

The main idea is to connect drop in R with drop in average mobility through basic linear regression.

R values were logarithmized in order to accurately present the rising difficulty of lowering R when it decreases (e.g. it is easier to lower R from 3 to 2 than it is from 2 to 1).

Because R on specific day represents the reproduction rate of virus on a given day, it is highly dependent on mobility during infectious period. That's why we decided to use future moving average for the following 8 days as the best mobility estimation for current day.

Regression results overview

Approach with log2 R and moving averaged mobility

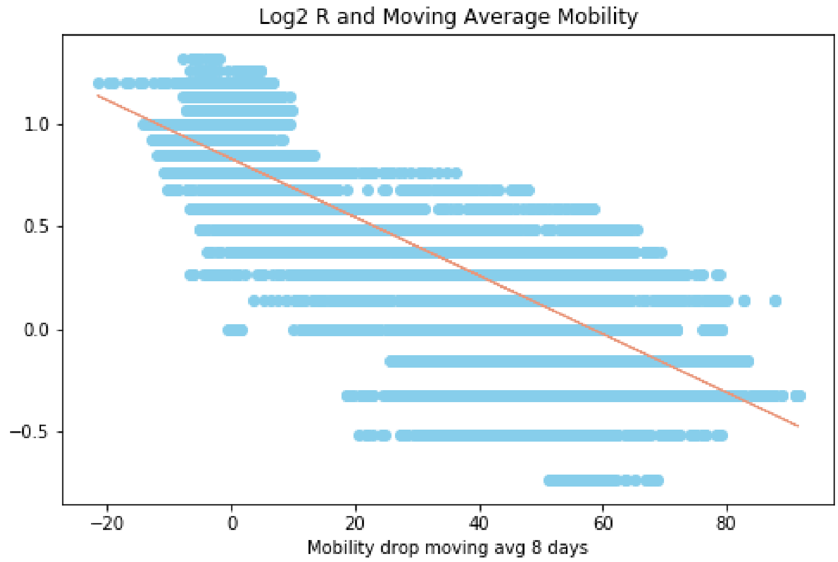


Figure 3: Linear regression for log2(R) and mov avg mobility

R^2: 0.676

Model P-value: 0.00

	value	p-value	95% conf interval
constant	0.8298	0.000	[0.8250, 0.835]
slope	-0.01426	0.000	[-0.014, -0.014]

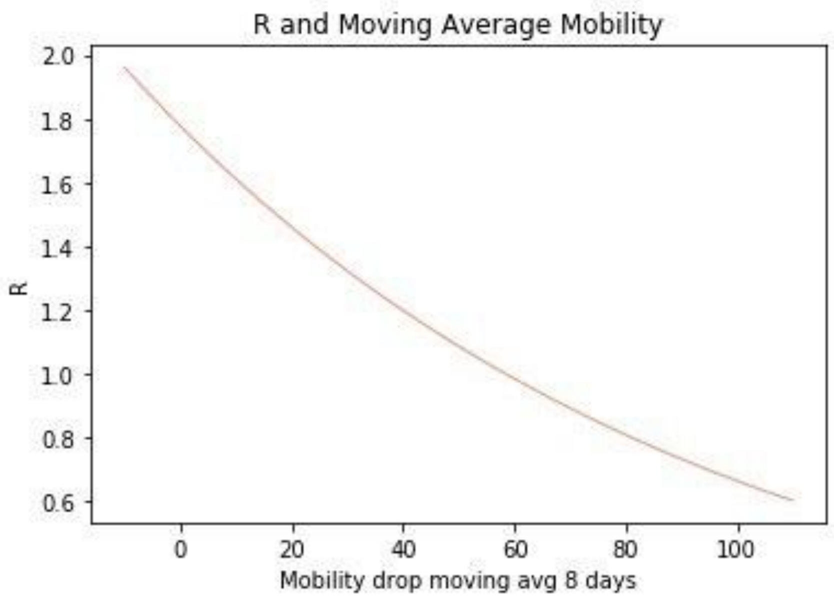


Figure 4: R as a function of mobility drop

As can be seen from Figure 4, in this model R decreases more rapidly from starting point than later. E.g. 0% to 20% mobility drop decreases R from 1.77 to 1.45 (difference of 0.32 in R) while 60% to 80% mobility drop decreases R from 0.98 to 0.81 (difference of 0.17 in R).