Draft Report

Team 7

If everyone could put their draft reports in their corresponding sections by Friday 20th Nov Morning for the meeting, that would be great. This document is meant to be the first draft of many and we be by no means complete.

Introduction (Everyone)

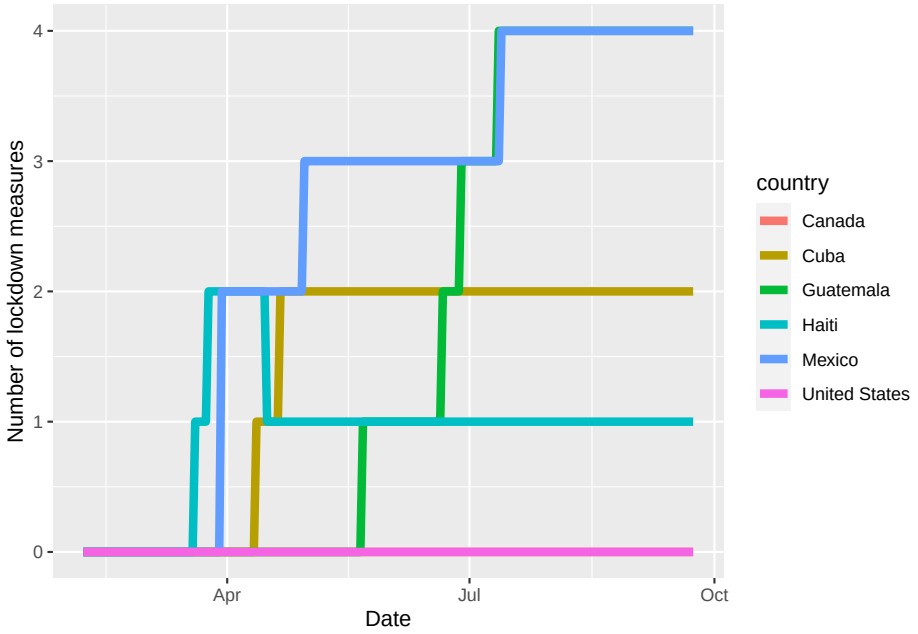
Section 1: Overview of lockdown (Vincent C)

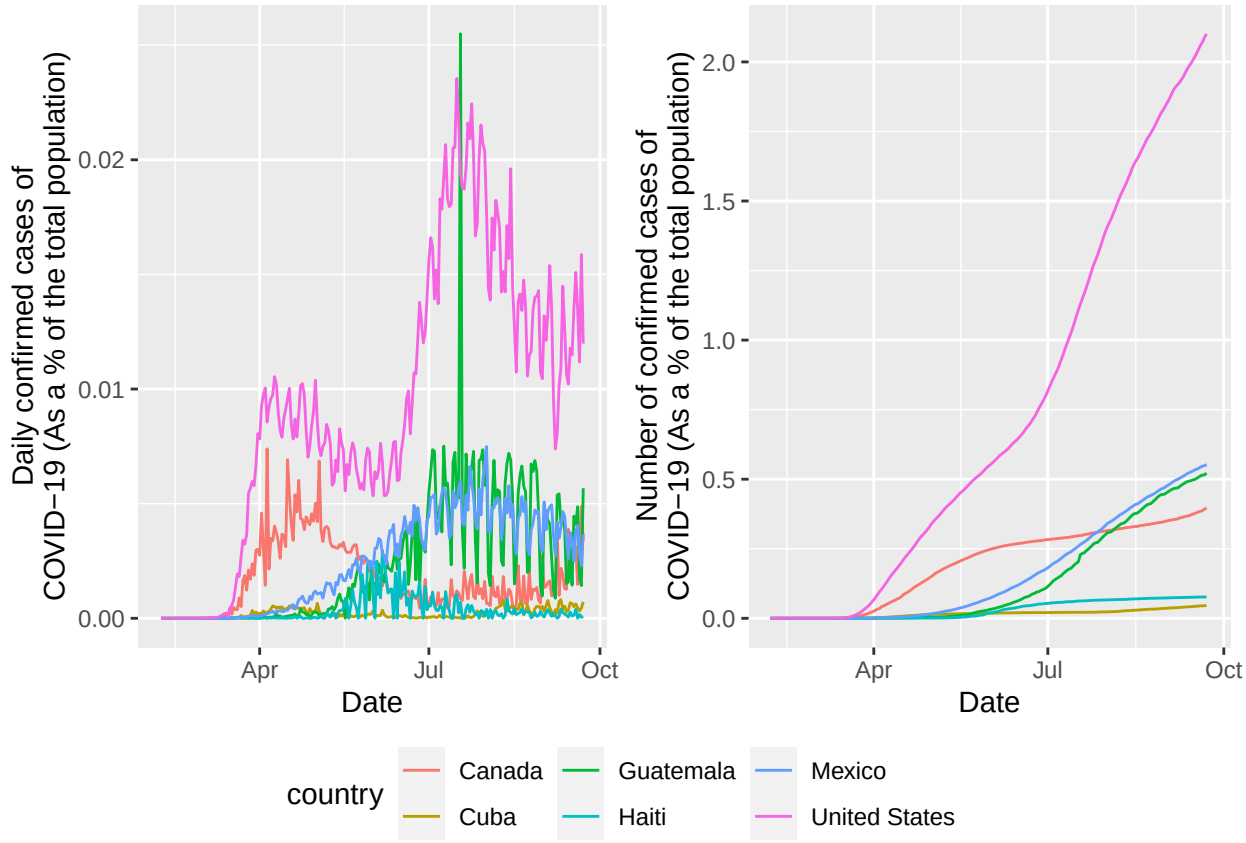
Overview of the spread of COVID-19

We will have a brief look into COVID-19 and its progression over the regions Europe, North America, Central America and South America. We will demonstrate how this virus escalated among the top 6 largest countries (by population) for each region by analyzing a few variables. The key variables that I believe displays this the best are the number of daily confirmed infections (as a % of the total population), the number of confirmed infections (as a % of the total population), and the number of lockdown measures for each country.

The number of lockdown measures will give us a high level overview into the response from governments as a result of the COVID-19 pandemic.

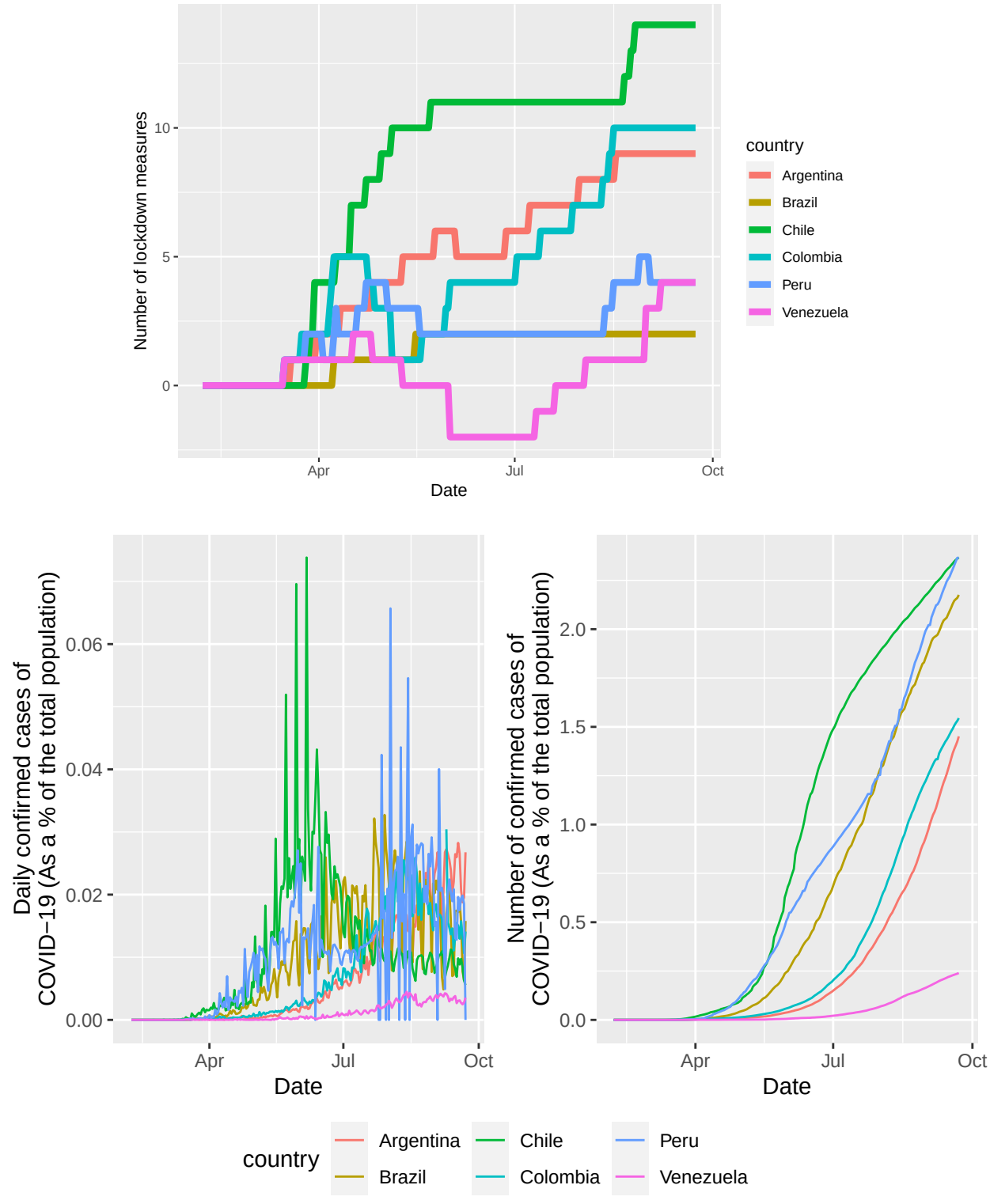
North America and Central America





From our plot, we can see that in the US and Canada the COVID-19 begun to take off around early April. However, for the other countries in this region the virus impacted them around May. We can see from the Lockdown Measures plot that the US and Canada had 0 measures in place to contain the virus, this must explain why they had a large number of cases early. We can also see that the other countries implemented lockdown measures around April, with Guatemala being the exception.

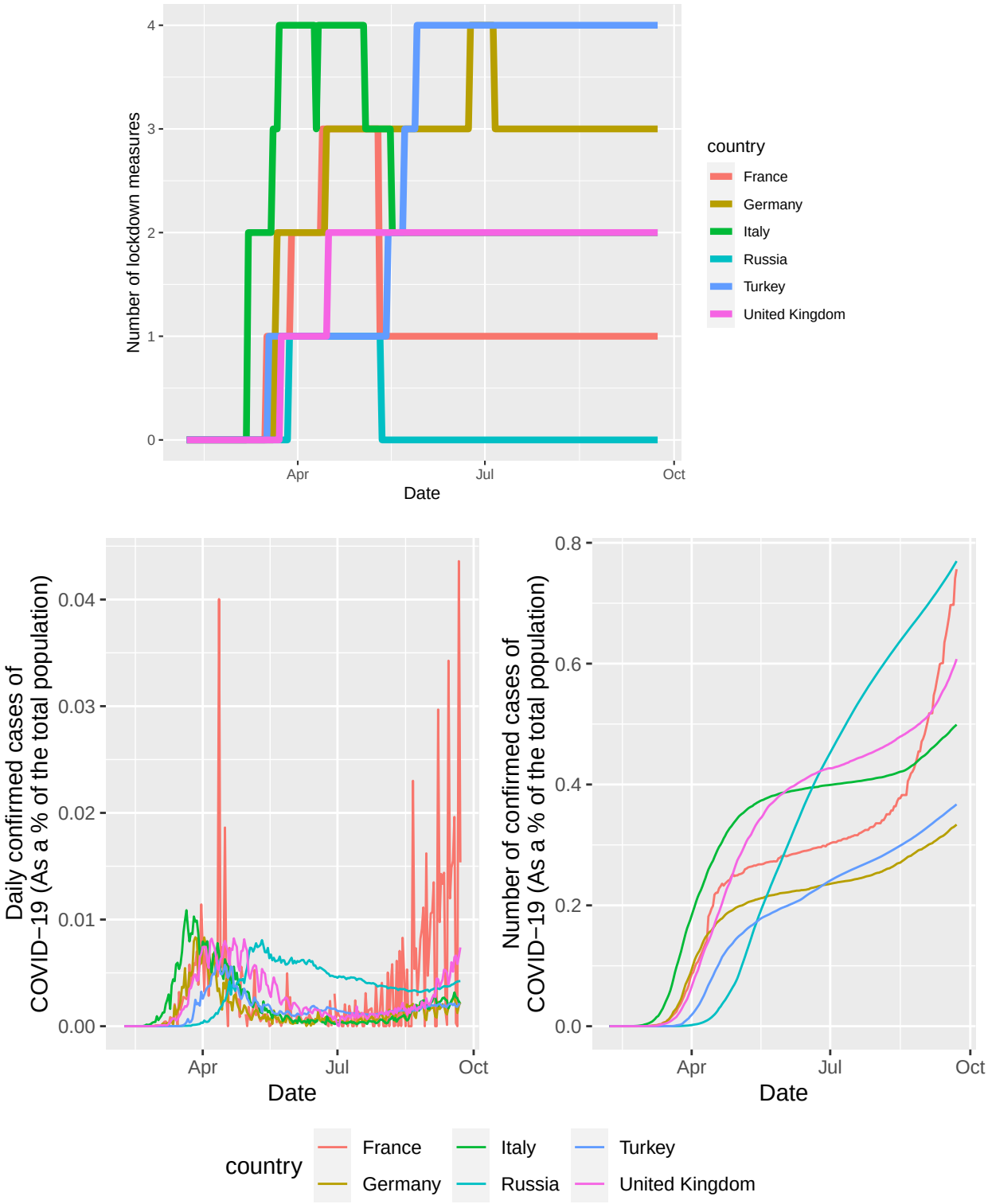
South America



From these set of graphs we can see that the the number of confirmed COVID-19 cases as a percentage of the population steadily rose later than in North and Central America. The cases started to rising around May. However, all the countries had at least one lockdown in early April, this may be one of the reasons why it impacted the countries in South America later than North and Central America. Looking at the daily

confirmed cases as a percentage of the population, we can see a growing increase of the infections in May. There is an unexpected result in the graph of the number of lockdown measures for Venezuela is showing a negative value between mid May to mid August. I could not find anything to explain this result.

Europe



Looking at the daily confirmed COVID-19 cases for Italy that the number of cases were on the rise in mid March. Other countries like France, Germany, United Kingdom and Turkey followed shortly afterwards in April. The spread of COVID-19 for Russia was roughly a month later than the other countries. Looking at the lockdown measures, by the beginning of April all the countries had at least one lockdown measure to help contain COVID-19. As we look at August we can see another rise in COVID-19 cases, this is where the second wave of COVID-19 occurred in Europe.

Section 2: Residential and workplace (Becky)

Introduction & data

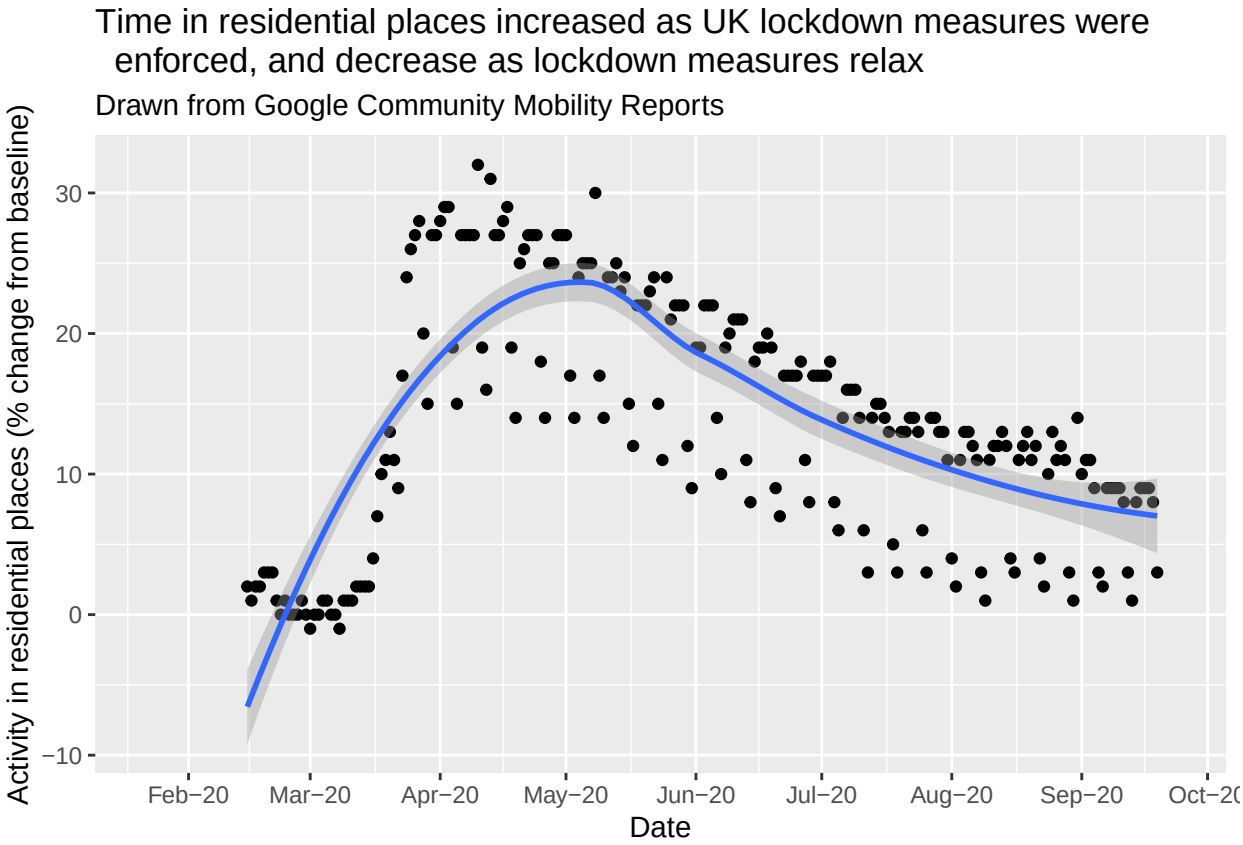
Since the large impacts of the pandemic on OmniCorp, particularly in the retail and hospitality sectors, we are interested in the effects of lockdown and other government interventions and how/why these might be different in different countries.

In the following, we will use the dataset "tidycovid19.csv", downloaded from the tidycovid19 R package on 24th September 2020. Descriptions of the different variables found in the data relating to the current epidemic and further details of the package can be found at this website (Gassen 2020).

We look at how the frequency of people visiting residential places and workplaces has changed during the pandemic for countries in Europe and America. In particular we study the `gcmr_residential` `gcmr_workplaces` variables from a community mobility report (Google, 2020). The variables are expressed as a percentage*100 change relative to the baseline period Jan 3 - Feb 6, 2020.

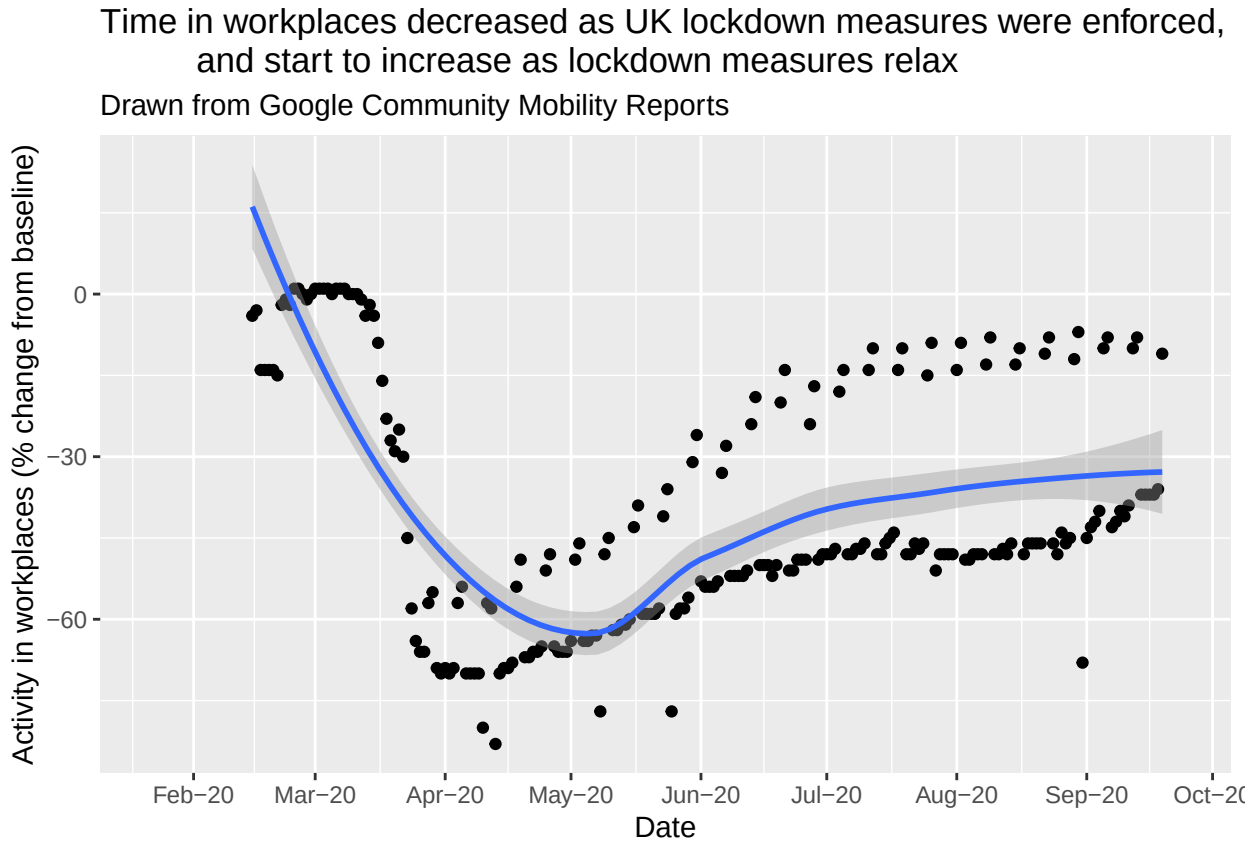
UK Data

We start by looking at how the frequency of visits to residential places has changed for UK. (need to delete one of graphs - wasn't sure if point or line graph better).



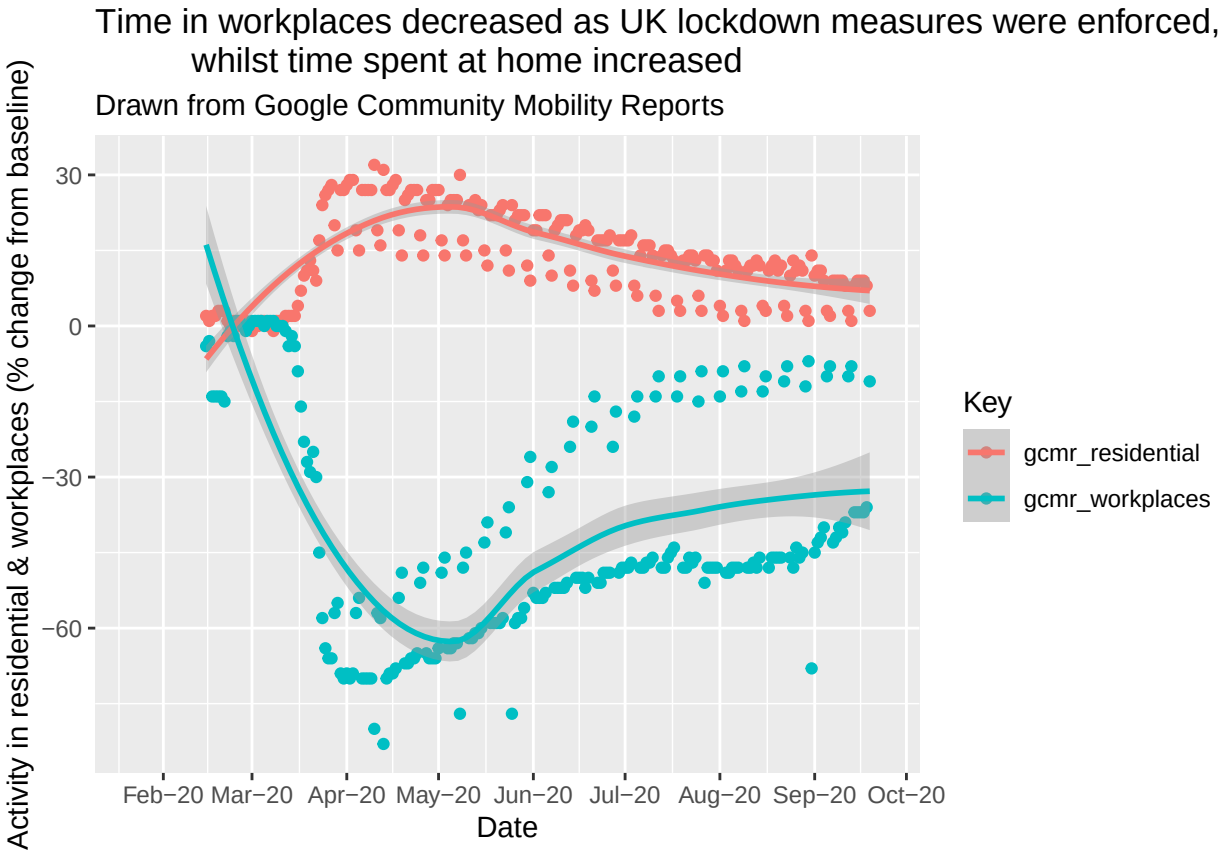
We note that there is an obvious increase in the time spent at home as lockdown is enforced on 24th March. In particular, the baseline measurement for activity in residential places is exceeded from this time onwards. However, we also see that the activity in residential places since is very volatile.

We now look at how the frequency of visits to workplaces has changed for UK. (need to delete one of graphs - wasn't sure if point or line graph better).



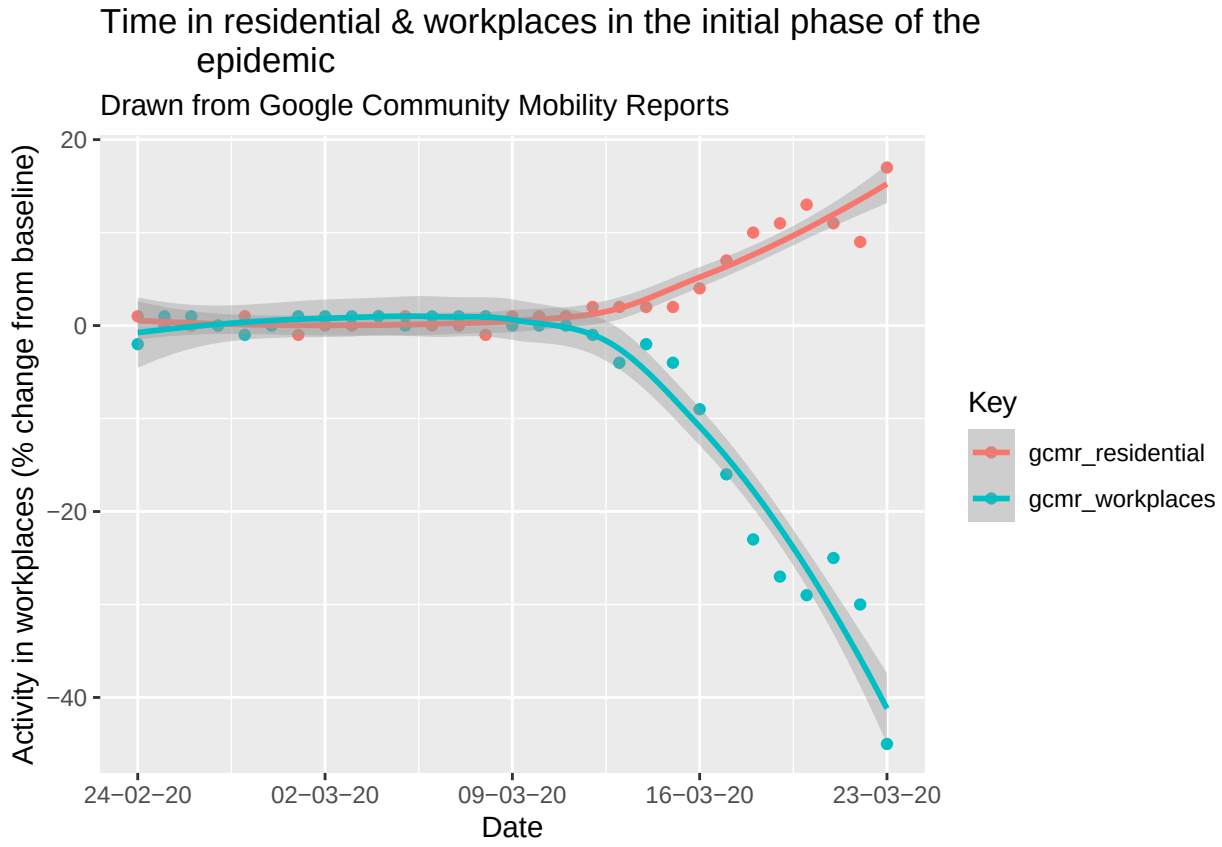
There is a dramtic decrease in the frequency of visits to workplaces since the UK lockdown is implemented on 24th March. The frequency of visits to workplaces does not recover to the baseline amount, even at the end of the time period it is far from it.

We see from these graphs that the `gcmr_residential` and `gcmr_workplaces` variables have an inverse relationship. We take a look at these variables for the UK on a single graph.



A plausible reason for this relationship is since lockdown measures were enforced by the government, everyone was asked to stay at home, and this included working from home. Therefore, a huge number of the working population were no longer travelling to work and instead remaining at home, and working remotely.

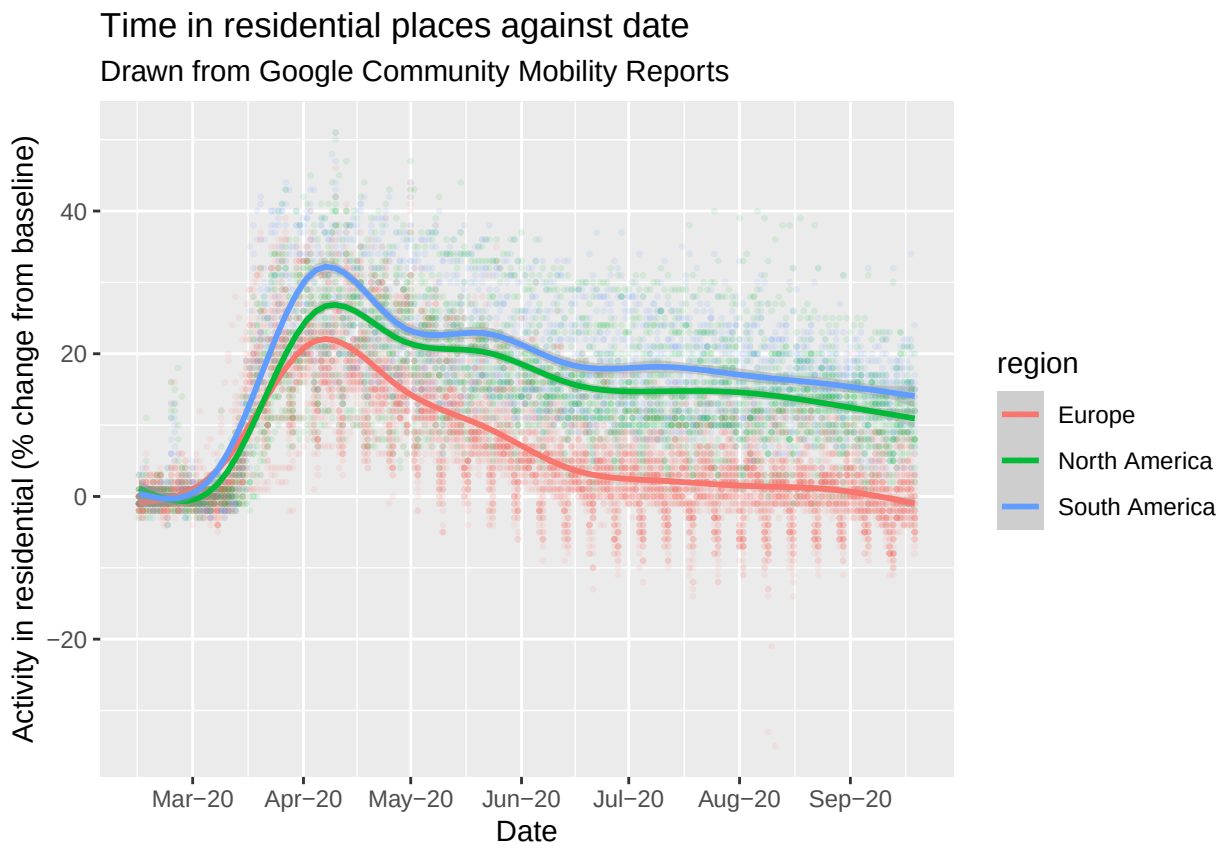
We look closer at the activity in residential & workplaces for UK in the initial phase of the pandemic. Since it is difficult to define precisely the 'first stage' of the pandemic in the UK, we look at the data from before the UK entered lockdown, 24th February - 23rd March.



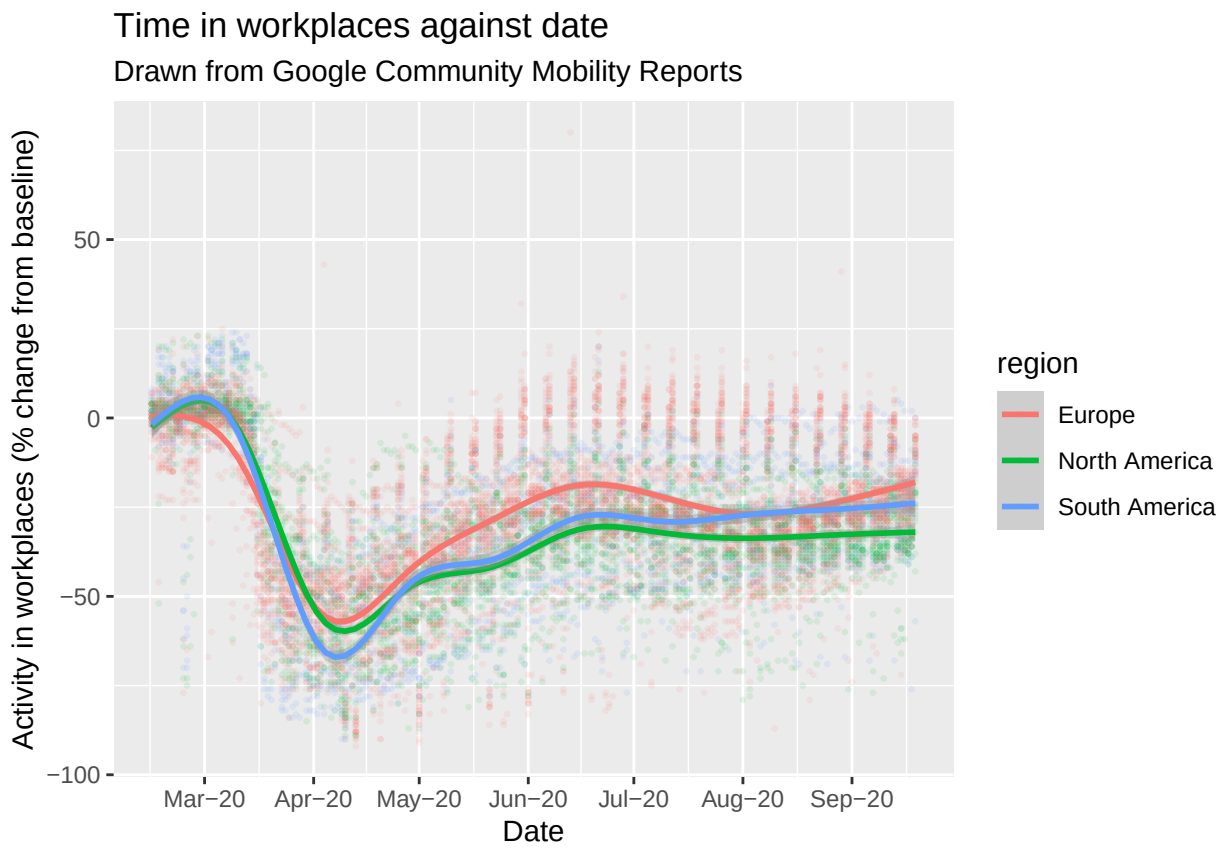
This plot confirms our thoughts of an inverse relationship between these variables.

Europe & America Data

We now look at the frequency of visits to residential places and workplaces for countries in Europe and America. We look particularly at the trends for Europe, North America and South America against date. We remove the data points and show the overall trend of the regions

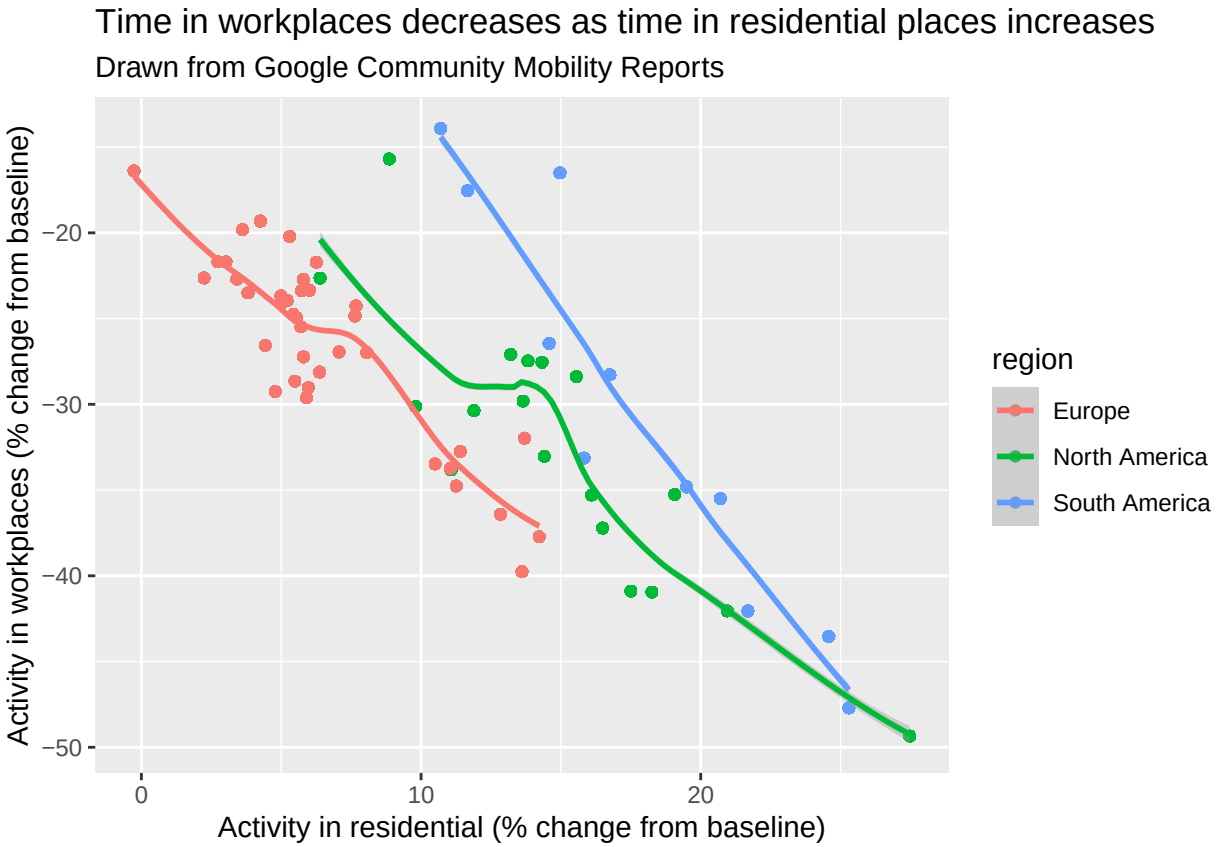


There is a huge increase in the time spent at home in all regions, with all peaks in April. The biggest spike is in South America, however this is potentially misleading since in this time period South America is in Autumn, in which people potentially stay at home more than in Spring.



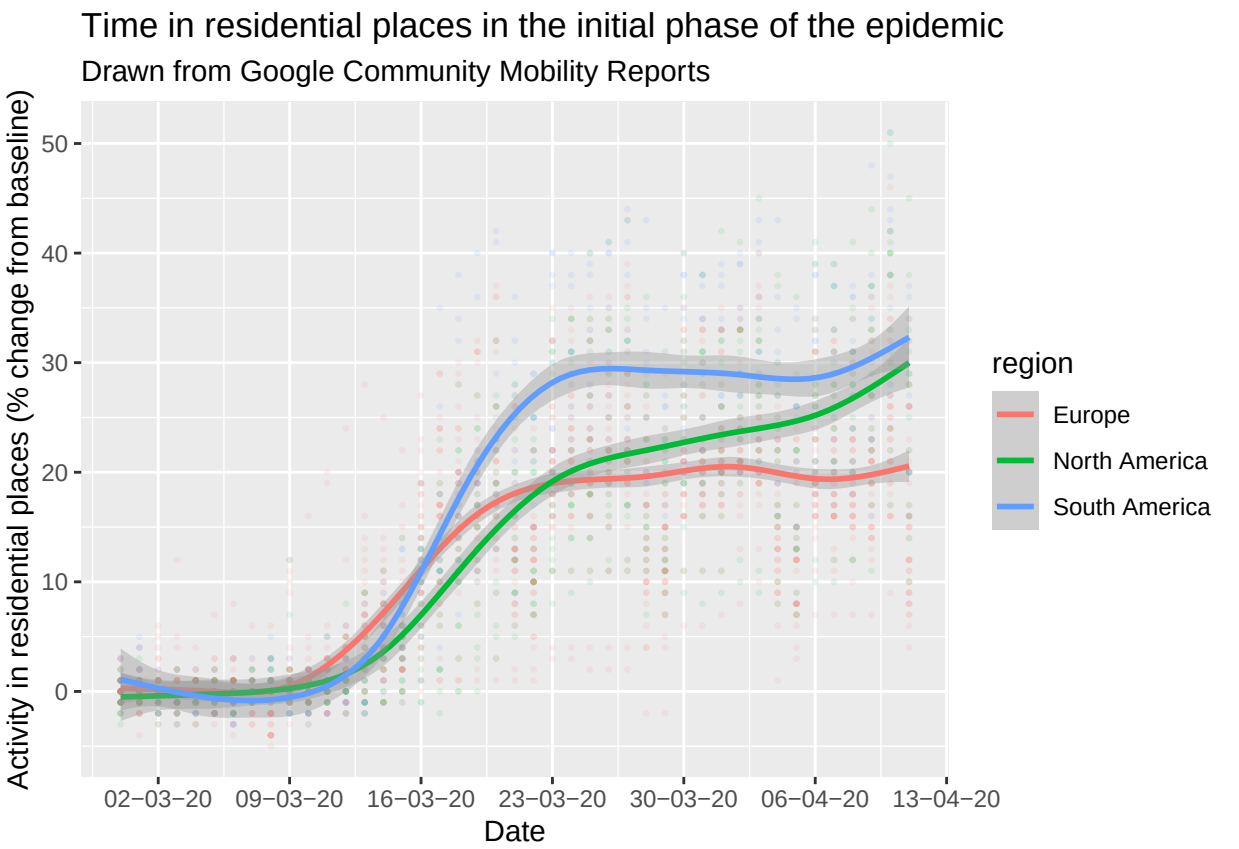
From these two plots we have the same inverse relationship between time in residential places and time in workplaces that we saw in the UK data.

We now look at the average percentage change in the frequency of visits to residential places against workplaces for countries in Europe and America from this year. We look particularly at the trends for Europe, North America and South America.

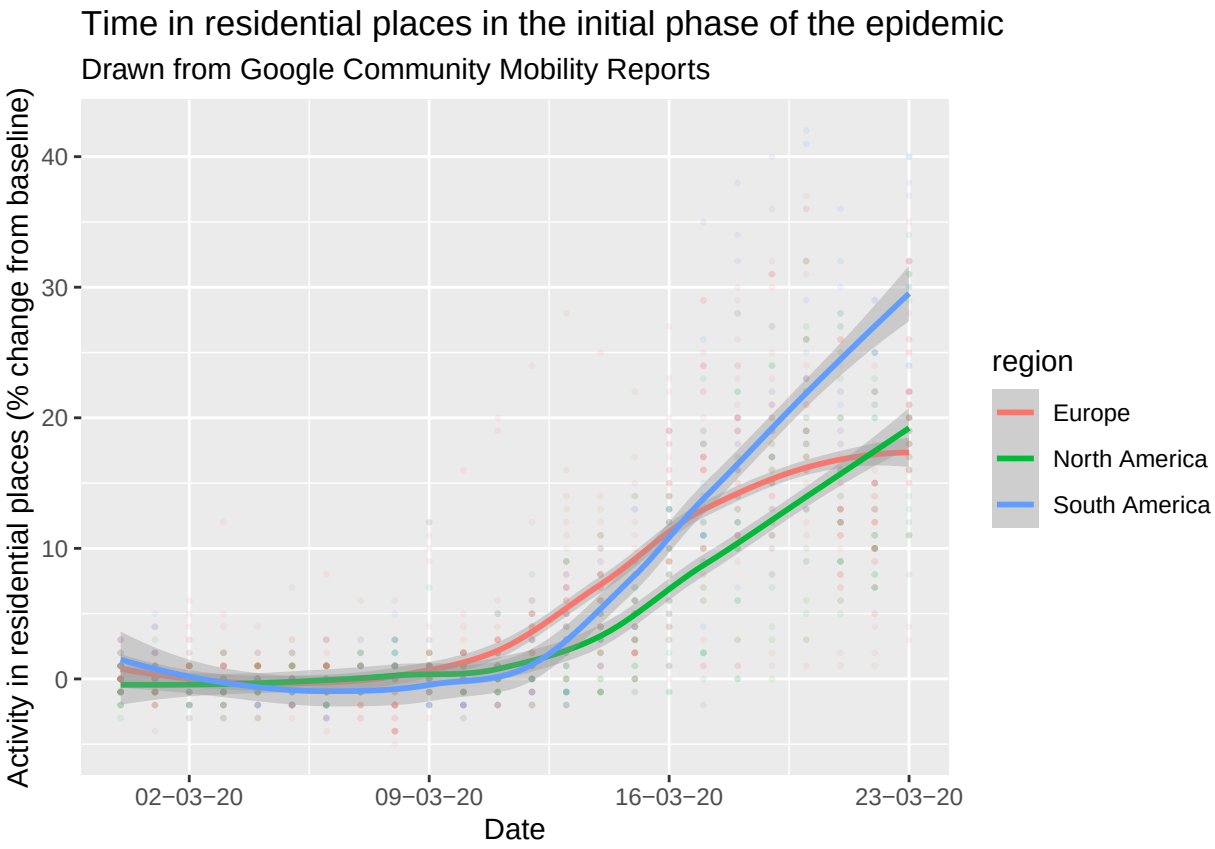


We see from this plot that as the time spent in workplaces decreases, the time spent in residential places increases, and this is true across the regions.

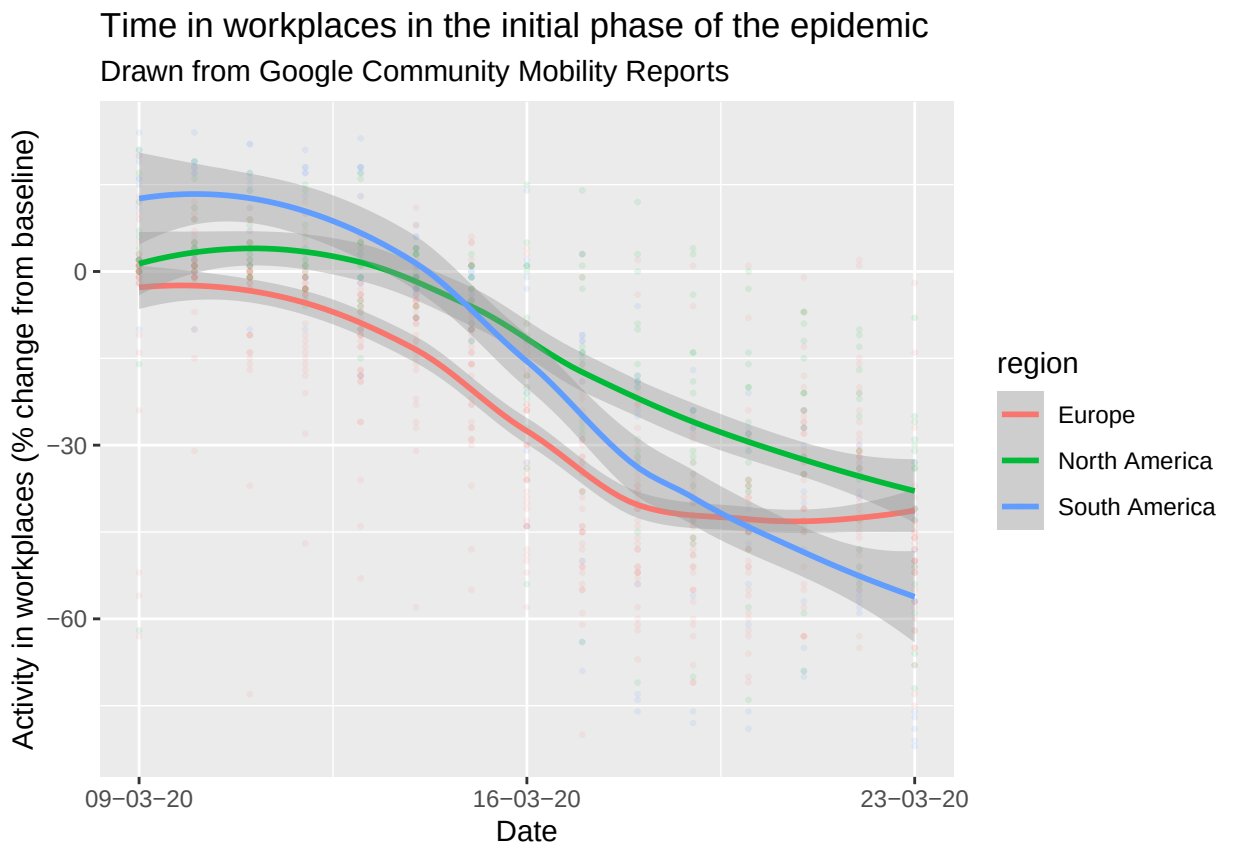
We look closer at the activity in homes and workplaces for these regions from 1st March - 11th Apr (note: these dates are subject to change from the data we gain from lockdowns). This is because we see a significant increase in activity in residential homes in this time period, as well as a significant decrease in activity in workplaces.



We look even closer at the data for time in residential places from before the UK entered lockdown, 1st - 23rd March. (Here I will link to lockdowns of diff countries, may change the time frame according, potential for modelling here)


Time in residential places in the initial phase of the epidemic
Drawn from Google Community Mobility Reports
Activity in residential places (% change from baseline)
40
30
20
10
0
02−03−20
09−03−20
16−03−20
23−03−20
Date
region
Europe
North America
South America

We see that from 9th - 23rd March approx. the time in workplaces falls in every region. We look closer at the data for this time period.



From this plot we see that both North & South America begin to increase and remain above the baseline reference, until about the 14th March, whilst Europe decreases throughout the time period, indicating that there were lockdown enforcements (such as working from home) earlier in Europe than in North & South America.

Conclusion

With more time spent at home and less time going into work there is less exposure to physical shops, i.e workers are no longer travelling past shops on their way into or home from work, or during their coffee/lunch breaks, so this results in an overall decrease in time spent in shops.

References

Gassen, 2020, *Download, Tidy and Visualize Covid-19 Related Data.* [online] Available at: https://joachim-gassen.github.io/tidycovid19/

Google, 2020. *COVID-19 Community Mobility Report.* [online] Available at: https://www.google.com/covid19/mobility/

Section 3: Grocery and pharmancy (Jacob)

To see how the trend of staying home has affected the retail and hospitality industry, we will start investigating consumer habits and how they changed throughout the pandemic.

To have a broader view of the impact of lockdowns on retail and hospitality, we need to examine the more fundamental parts of people's consumer habits. We first start by investigating people's visits to grocery shops and pharmacies. Although this sector isn't directly connected to retail and hospitality, the principles are the

same - they sell goods to consumers directly. To understand the willingness of people to go outside and buy clothes, we need to first understand their willingness to go to a grocery shop or pharmacy.

Below you can see a graph depicting average the percentage change in the frequency of people's visits to grocery stores and pharmacies, relative to a baseline period Jan3- Feb 6, 2020 of the three regions Omnicorp is operating in.

```
library(readxl)
CovidData <- read.csv("tidycovid19.csv")
```

```
library(lubridate)
CovidData$date = as.Date(parse_date_time(CovidData$date,orders=c("y","ym","ymd")))
```

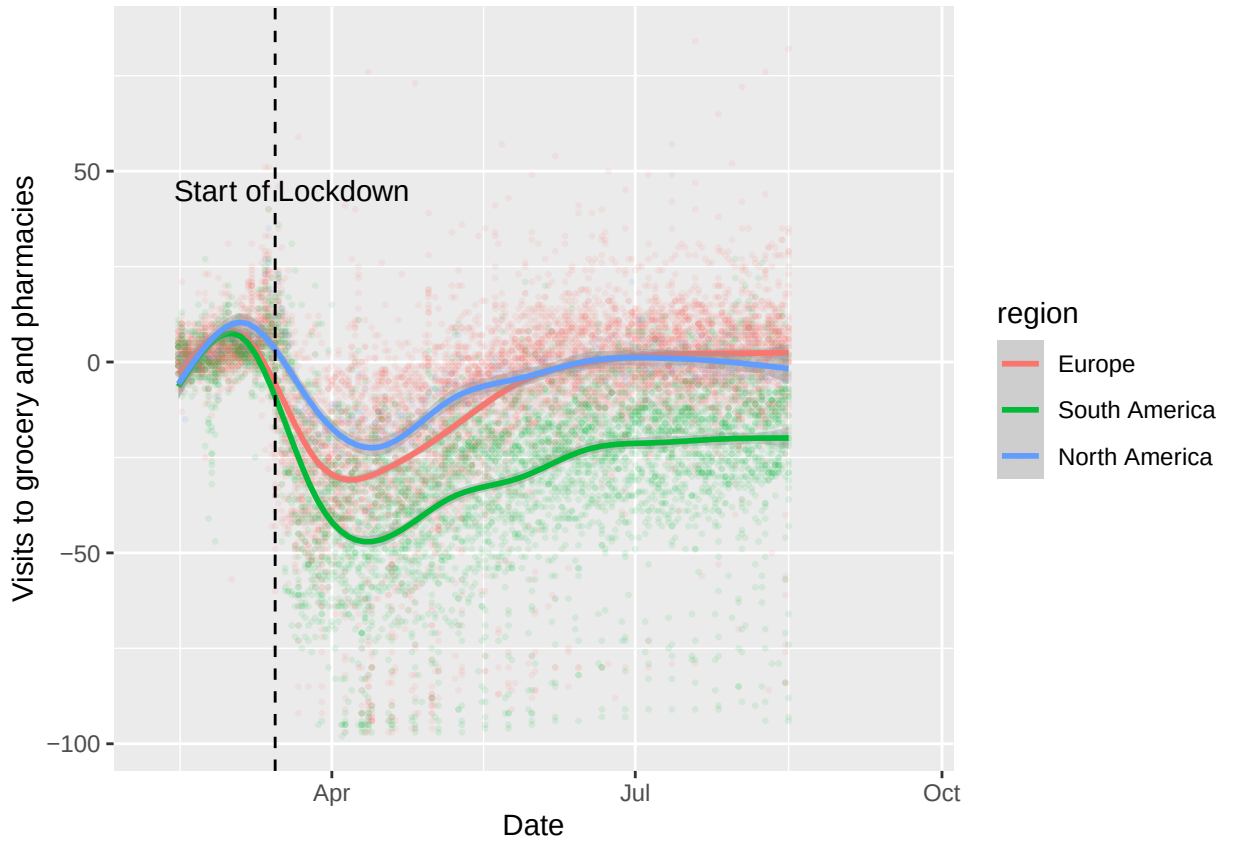
```
library(dplyr)
confirmedData <- {CovidData %>%
    select(date, confirmed, gcmr_grocery_pharmacy) }
```

```
library(stringr)
EuropeAmericaDataPostFeb6 <- {CovidData %>%
    filter( ( str_detect(region, "America") | str_detect(region, "Europe") ) & (date > as.Date("2020-02
```

```
groceryData = {EuropeAmericaDataPostFeb6 %>%
    group_by(country) %>%
    summarise(
              total_confirmed = max(confirmed),
              mean_gcmr_grocery_pharmacy = mean(gcmr_grocery_pharmacy),
              pop_largest_city = mean(pop_largest_city ))}
```

```
mean_ <- function(...) mean(..., na.rm=T)
max_ <- function(...) max(..., na.rm=T)
groceryData = {EuropeAmericaDataPostFeb6 %>%
    group_by(country) %>%
    summarise(
              total_confirmed = max_(confirmed),
              mean_gcmr_grocery_pharmacy = mean_(gcmr_grocery_pharmacy),
              pop_largest_city = mean(pop_largest_city ))}
```

```
library(ggplot2)
EuropeAmericaDataPostFeb6_copy=EuropeAmericaDataPostFeb6
plot1 <- {EuropeAmericaDataPostFeb6_copy %>% ggplot(aes(x=date,
                                        y=gcmr_grocery_pharmacy,color=region))} +
    geom_point(alpha=0.1,size=0.5)+geom_smooth()+scale_colour_discrete(labels = c('Europe', 'South Amer
print(plot1)
```



We see a sharp decline in visits starting in March and continuing until the start of September. South America has been affected more heavily than the others, as there have been more strict and continuous lockdown measures, as seen in figure #. Although we see South America recovers slower, we should take into consideration the fact that during these months, the southern hemisphere is in Winter, and during winter people tend to stay home more.

Section 4: Recreation and retail (Vincent P)

To help Omnicorp to get a general picture on how pandemic impacted the whole retailing and recreation industries, we now investigate the change of activities of retailing and recreational places people visited from a baseline level.

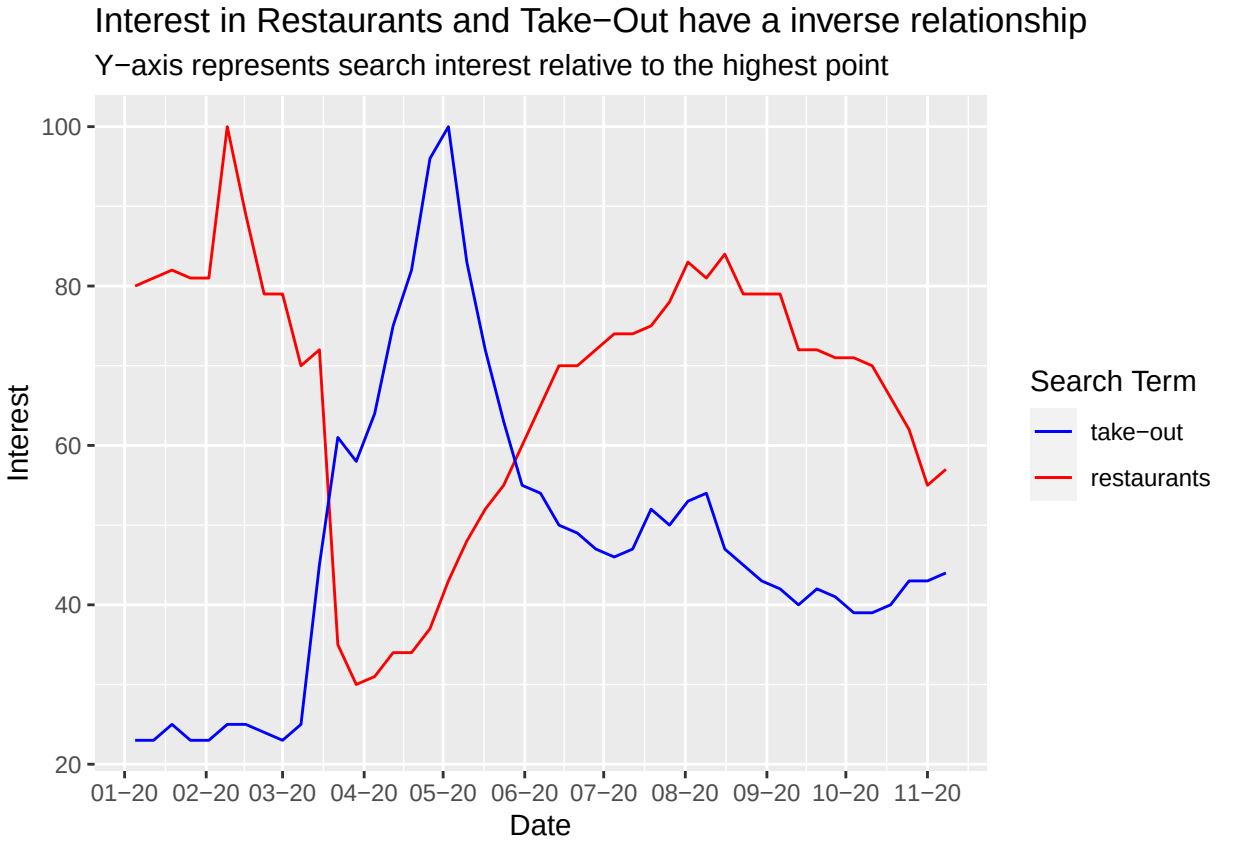
Furthermore, we are not only interested in the offline retailing data collected by COVID19??(ensure later), we would like to know how online business activities changed. Yet first, let's see the development of internet users in the three continents.

Section 5: Google Trends (Max)

Using Google Trends data on the search volume for various terms, we can investigate the general interest over time. Google Trends is an unbiased sample of Google search data. It's anonymised, categorized and aggregated. For the regions we are looking at - Europe, North, Central and South America - the percentage of population that uses the internet is 88%, 95%, 61% and 72% respectively. This indicates that the search patterns shown by Google Trends may be an accurate representation of the behaviors and interests of these regions.

We will investigate the search terms `restaurants` and `take-out`, the two can be seen as the compliment of one another. We will plot a graph of the average interest for these terms throughout European, North,

Central and South American countries. The y axis represents search interest relative to the highest point on the chart for the given region and time. A value of 100 is the peak popularity for the term. A value of 50 means that the term is half as popular.



The plot above shows that the two search terms' interest have an inverse relationship. As we can see, when lockdown and restrictions began to be imposed by governments, consumer interest in `restaurants` decreased whereas interest in `take-out` increased at a similar magnitude. This shows that in lockdowns, consumers are looking to buy more food online, possibly due to sit-in restaurants being closed or people scared to go out in public. This indicates that we, OmniCorp, should focus on online retail and hospitality when lockdown measures are introduced. It is also important to note that the interest in `restaurants` increased back to normal levels almost as fast as it fell. However, it had a smaller, more steady decrease as a second wave of restrictions took place whereas `takeout` did not have a similar magnitude of increase. # Conclusion (Everyone)