Abstract

Our goal is to provide some summary statistics of deaths across countries during the 1st Wave of Covid-19 and to compare these numbers with the corresponding ones of the previous years. This analysis is not scientific and we cannot drive any conclusion about the impact of Covid-19 since we need to take into consideration many other parameters and this is out of the scope of this analysis.

We still keep our character which is a Data Science focusing on R and Python and for that reason, we will work with R and we will share the code making the analysis to be reproducible.

Data Source

We found many difficulties to find structured, reliable and consistent data about deaths across countries. For this analysis we obtained the data from the Our World in Data and more particularly at the Excess Mortality using raw deaths counts and then we downloaded the excess-mortality-raw-death-counts.csv file.

Some notes about the data:

- The number of deaths is measured by week. So, if we decide to report the results by month then some months will have more weeks than the others, that is why is better to report the average weekly deaths by months instead of total.
- Week dates are defined by international standard ISO 8601, although not all countries follow this standard.
- The Human Mortality Database has data for England & Wales (and Scotland) but not for the UK as a whole. England & Wales compose ~89% of the UK population. Source: UK ONS Population estimates for the UK, England and Wales, Scotland and Northern Ireland: mid-2019
- For some countries, there exist data up to September 2020 wherein some other up to October 2020. It takes some weeks for the database to be updated.
- For this analysis, we will include countries of the Northern Hemisphere and those with a high number of Corona Virus Cases per Population like the **US**, **Italy**, **Spain**, **France** etc.

Summary Statistics by Country

The data is from 2015 up to 2020 and is referring to the same period (for example January to September). We will provide the following summary statistics by country:

- A table of Average Weekly Deaths by Month
- A chart of Average Weekly Deaths by Month
- Total Deaths per Year for the same period
- Percentage difference in Total deaths of **2020** compared to **2015, 2016, 2017, 2018 and 2019**

Prepare the data:

```
library(tidyverse)
```

```
# https://ourworldindata.org/excess-mortality-covid?fbclid=IwAR1LZGFQGI_
nXmVMMnhnXOkLzhYTHglGPGYBIKyxOQCP3L2HLr9H99XNgE0
df<-read.csv("excess-mortality-raw-death-count.csv")

# Rename the column names
names(df)<-c("Country", "Code", "Date", "Deaths_2020", "Deaths_AVG_2015_2019"
             "Deaths_2015", "Deaths_2016", "Deaths_2017", "Deaths_2018",
"Deaths_2019")



# Convert the Date to Date-Format
df$Date<-as.Date(df$Date)
```

US

Let's provide the summary statistics of the US.

```
# Excess Deaths by Month in US
# Create also a Date Month column where we truncate the date to the month
# Create also a month column which is the month
country<- df%>%filter(Code=="USA")%>%
  mutate(Date_Month=lubridate::floor_date(Date, unit = "month"), Month =
lubridate::month(Date, label=TRUE))



country%>%select(-Deaths_AVG_2015_2019)%>%group_by(Month)%>%
  summarise(across(Deaths_2020:Deaths_2019, ~mean(.x, na.rm=TRUE)))
```

Table of Average Weekly Deaths by Month

Month	Deaths_2020	Deaths_2015	Deaths_2016	Deaths_2017	Deaths_2018	Deaths_2019
<ord>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1 Jan	59548.	60206.	55522.	59516.	65260.	58164
2 Feb	58726	56824	55159.	58555.	59911.	58096.
3 Mar	59738.	54825.	54994.	56821.	56153.	57612.
4 Apr	75198.	52533.	52927.	54092.	54564.	55042.
5 May	64102.	50562.	50582.	51865.	51682.	53471.
6 Jun	58002	49232.	49795	51001.	51319.	52550
7 Jul	61658.	48968.	49571.	50102.	50935	51867.
8 Aug	60999	48175.	49623.	50168.	50589.	51203.
9 Sep	55514	48714	50066	50973	50292	51772

Chart of Average Weekly Deaths by Month

```
country%>%select(-Deaths_AVG_2015_2019)%>%group_by(Month)%>%
  summarise(across(Deaths_2020:Deaths_2019, ~mean(.x, na.rm=TRUE)))%>%
  pivot_longer(-Month, names_to="Year", values_to="Total_Deaths")%>%
  ggplot(aes(x=Month, y=Total_Deaths, group=Year))+
```

```
geom_line(aes(col=Year))+geom_point(aes(col=Year))+
ggtitle("US: Avg Weekly Deaths by Month")+ylab("Avg Weekly Deaths within
Month")+
theme_minimal()
```
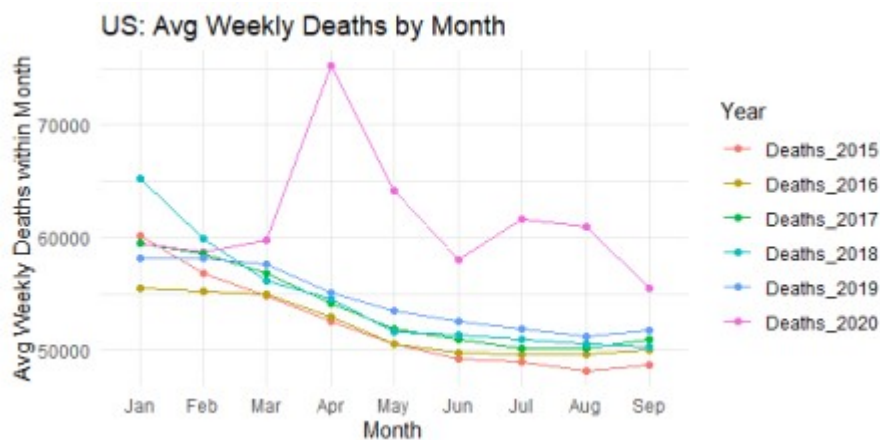

As we can see, in **2020** the deaths during January and February were very closed with the past years and then, from February onwards there is a significant increase.

Total Deaths per Year for the same period

```
country%>%select(-Deaths_AVG_2015_2019)%>%
summarise(across(Deaths_2020:Deaths_2019, ~sum(.x, na.rm=TRUE)))
```

	Deaths_2020	Deaths_2015	Deaths_2016	Deaths_2017	Deaths_2018	Deaths_2019
1	2232237	1887579	1877956	1938312	1970369	1966086

Percentage difference in Total deaths of 2020 compared to previous years

```
# % Increase in deaths
country%>%select(-Deaths_AVG_2015_2019)%>%
summarise(sum(Deaths_2020)/across(Deaths_2020:Deaths_2019, ~sum(.x,
na.rm=TRUE))-1)
```
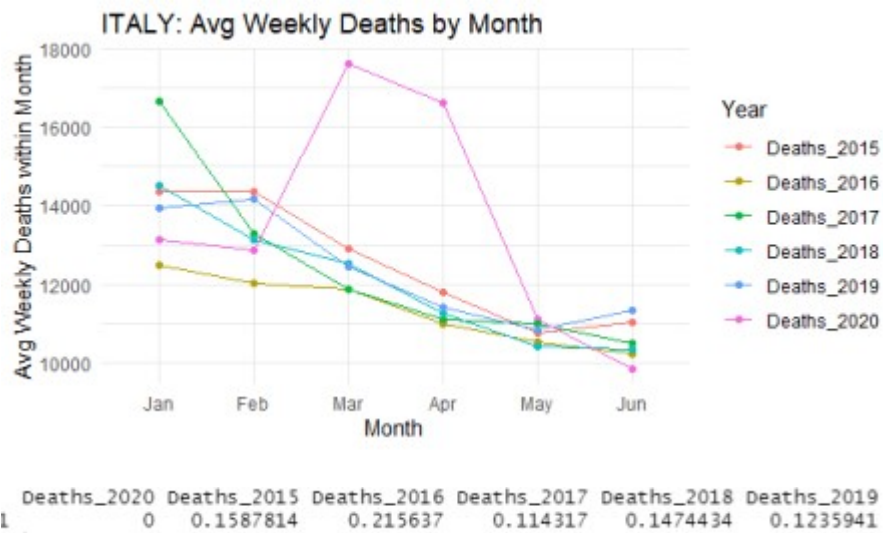
	Deaths_2020	Deaths_2015	Deaths_2016	Deaths_2017	Deaths_2018	Deaths_2019
1	0	0.1825926	0.1886524	0.1516397	0.132903	0.135371

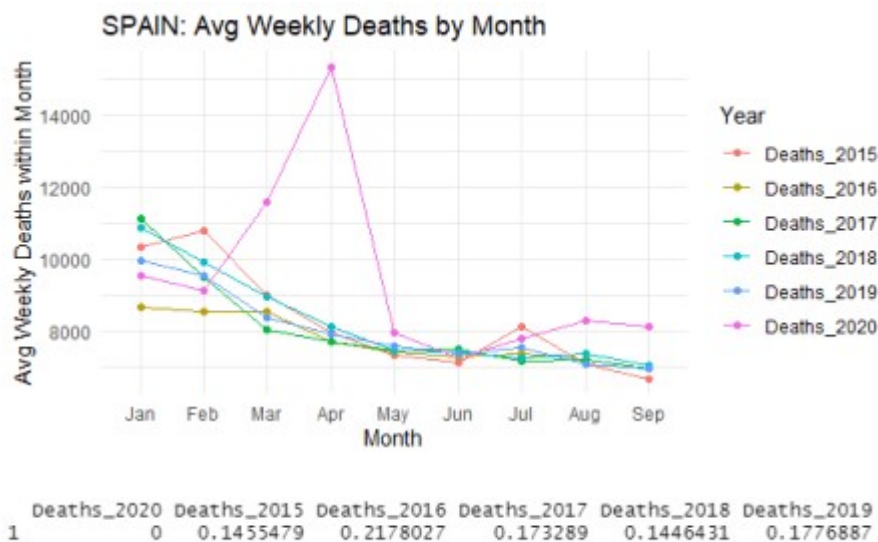We can see that compared to 2019, **there is an increase of 13.53% in deaths**.

Summary Statistics for Other Countries

We provided a detailed example of how you can generate these statistics for a specific country. You can simply change the country code. Let's provide the charts and the percentage of increased deaths compared to previous years for the other countries.
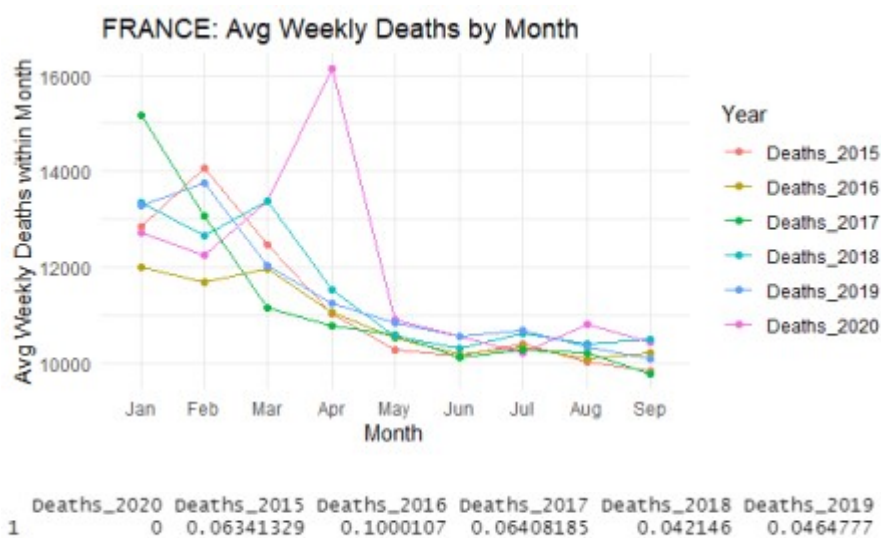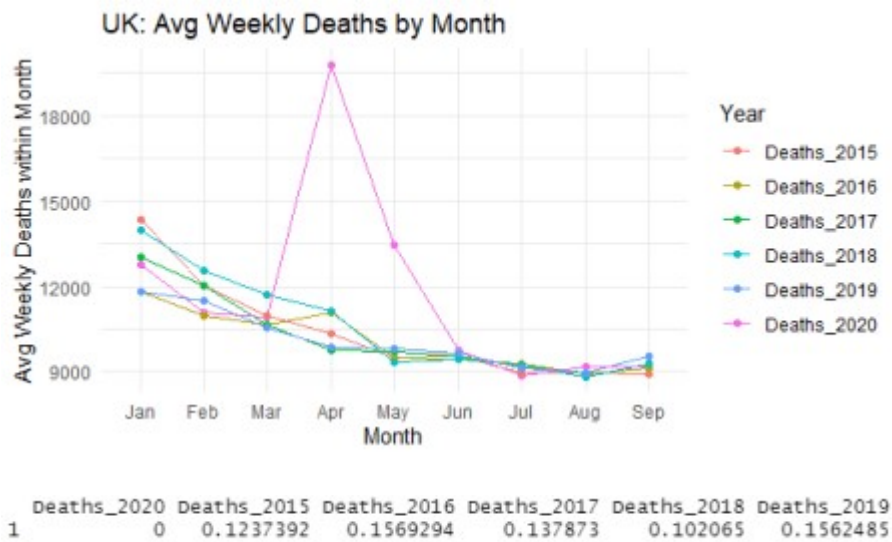
Italy

ITALY: Avg Weekly Deaths by Month

	Deaths_2020	Deaths_2015	Deaths_2016	Deaths_2017	Deaths_2018	Deaths_2019
1	0	0.1587814	0.215637	0.114317	0.1474434	0.1235941

Spain

SPAIN: Avg Weekly Deaths by Month

	Deaths_2020	Deaths_2015	Deaths_2016	Deaths_2017	Deaths_2018	Deaths_2019
1	0	0.1455479	0.2178027	0.173289	0.1446431	0.1776887

France

FRANCE: Avg Weekly Deaths by Month

	Deaths_2020	Deaths_2015	Deaths_2016	Deaths_2017	Deaths_2018	Deaths_2019
1	0	0.06341329	0.1000107	0.06408185	0.042146	0.0464777

England and Wales

UK: Avg Weekly Deaths by Month

	Deaths_2020	Deaths_2015	Deaths_2016	Deaths_2017	Deaths_2018	Deaths_2019
1	0	0.1237392	0.1569294	0.137873	0.102065	0.1562485

Sweden

SWEDEN: Avg Weekly Deaths by Month

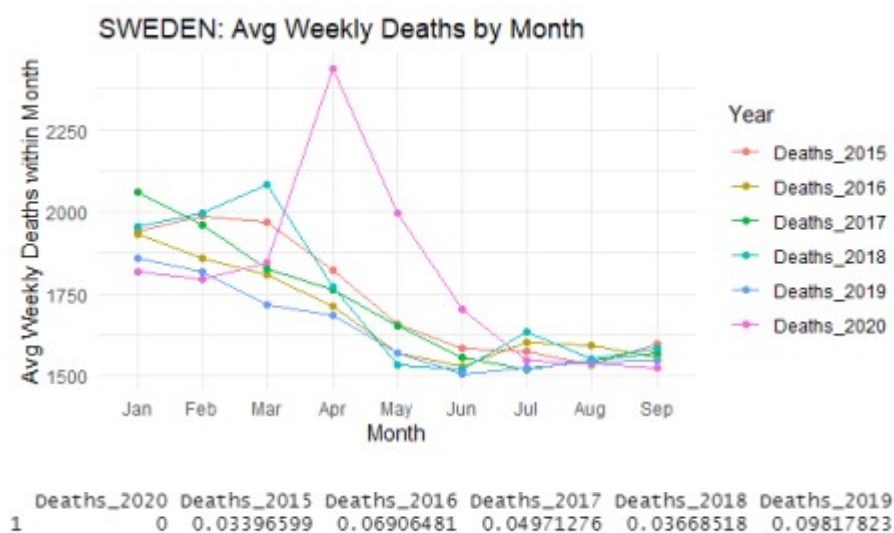|	Deaths_2020	Deaths_2015	Deaths_2016	Deaths_2017	Deaths_2018	Deaths_2019
1	0	0.03396599	0.06906481	0.04971276	0.03668518	0.09817823

Discussion

From the data above we may argue that for the countries that have many Covid-19 cases, there is a significant increase in raw deaths during the March-May period and then during the summer period the raw deaths for most the countries revert back to its "normal" levels. All of these countries above showed an increase in deaths compared to the previous years. For example, Spain has **17.76%** more deaths compared to **2019** for the same period (Jan-Sep). The corresponding numbers for the rest countries are **US 13.53%, Italy 12.35%, France 4.64%, England and Wales 15.62% and Sweden 9.81%**.

Again we want to stress out that we cannot guarantee the validity of our data and we do not take into consideration other parameters. For example, The pandemic may result in fewer deaths from other causes – for instance, the mobility restrictions during the pandemic might lead to fewer deaths from road accidents. From the data above it seems that there is an "anomaly" in deaths during the 1st Wave of Covid-19. If this Covid-19 is a hoax according to our friends below, then what is this external factor which causes more deaths?