

# Quelques chiffres autour de l'épidémie de Covid-19

François Févotte

April 19, 2020

## Contents

<b>1</b>	<b>Identification du système</b>	<b>1</b>
<b>2</b>	<b>Récupération des données</b>	<b>1</b>
<b>3</b>	<b>Evolution comparée de l'épidémie dans quelques pays</b>	<b>2</b>
<b>4</b>	<b>Vers un passage du pic ?</b>	<b>4</b>

## 1 Identification du système

```
date "+%Y-%m-%d"
```

```
2020-04-19
```

```
uname -a
```

```
Linux aluminium 4.14.0-3-amd64 #1 SMP Debian 4.14.17-1 (2018-02-14) x86_64 GNU/Linux
```

```
import sys
sys.version
```

```
3.7.3rc1 (default, Mar 13 2019, 11:01:15)
[GCC 8.3.0]
```

## 2 Récupération des données

On prend les données distribuées par Our World In Data, elles-mêmes issues de l'ECDC :

```
url = "https://covid.ourworldindata.org/data/ecdc/full_data.csv"
data_raw = pd.read_csv(url, index_col='date')
data_raw.tail(3)
```

	location	new_cases	new_deaths	total_cases	total_deaths
date					
2020-04-16	Zimbabwe	6	0	23	3
2020-04-17	Zimbabwe	1	0	24	3
2020-04-18	Zimbabwe	0	0	24	3

### 3 Evolution comparée de l'épidémie dans quelques pays

Je préfère fonder l'analyse sur le nombre de décès, qui me semble plus fiable et comparable d'un pays à l'autre que, par exemple, le nombre de cas dépistés. Afin de recaler l'axe temporel, on positionne  $t = 0$  au moment où le nombre total de morts dépasse 100.

La logique est la suivante : on considère le nombre  $n^c$  (de décès ou de cas) dans un pays  $c$ . On suppose que l'évolution de  $n^c$  en fonction du temps  $t$  suit une loi de croissance exponentielle (au moins dans une phase initiale). On s'attend ainsi à avoir une expression du type :

$$n^c(t) = n_0^c e^{\lambda^c(t-t_0^c)}.$$

Dans cette expression, la valeur initiale  $n_0^c = n^c(t_0^c)$  est liée à la taille de la population considérée. La constante de temps  $\lambda^c$  en revanche, n'a pas vraiment de raison de l'être : elle reflète simplement le rythme auquel l'épidémie se propage.

Si l'on se donne un seuil  $\tau$  unique, on peut fixer pour chaque pays l'instant initial  $t_0^c$  de sorte que

$$n^c(t_0^c) = \tau.$$

En effectuant un recalage de l'instant initial, c'est à dire en effectuant le changement de variable temporelle  $\tilde{t}^c = t - t_0^c$ , on obtient

$$n^c(\tilde{t}^c) = \tau e^{\lambda^c \tilde{t}^c},$$

ce qui signifie que tous les pays deviennent comparables, indépendamment de leur taille : on s'est ramené à comparer les constantes de temps  $\lambda^c$  entre elles.

En pratique, on prend

$$t_0^c = \min \{t; n^c(t) > \tau\}.$$

```

# Parameters
column      = "total_deaths"
threshold   = 100
countries   = ["Italy", "Spain", "France", "United Kingdom", "Germany", "Switzerland"]

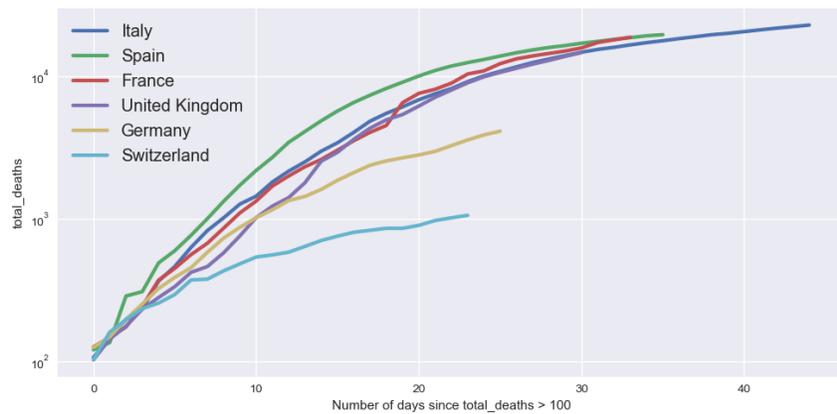
total = {}
for country in countries:
    total[country] = data_raw.query(f"location == '{country}' & {column} > {threshold}")

total[countries[2]]

```

date	location	new_cases	new_deaths	total_cases	total_deaths
2020-03-16	France	924	36	5423	127
2020-03-17	France	1210	21	6633	148
2020-03-18	France	1097	27	7730	175
2020-03-19	France	1404	69	9134	244
2020-03-20	France	1861	128	10995	372
2020-03-21	France	1617	78	12612	450
2020-03-22	France	1847	112	14459	562
2020-03-23	France	1559	112	16018	674
2020-03-24	France	3838	186	19856	860
2020-03-25	France	2446	240	22302	1100
2020-03-26	France	2931	231	25233	1331
2020-03-27	France	3922	365	29155	1696
2020-03-28	France	3809	299	32964	1995
2020-03-29	France	4611	319	37575	2314
2020-03-30	France	2599	292	40174	2606
2020-03-31	France	4376	418	44550	3024
2020-04-01	France	7578	499	52128	3523
2020-04-02	France	4861	509	56989	4032
2020-04-03	France	2116	471	59105	4503
2020-04-04	France	5233	2004	64338	6507
2020-04-05	France	4267	1053	68605	7560
2020-04-06	France	1873	518	70478	8078
2020-04-07	France	3912	833	74390	8911
2020-04-08	France	3777	1417	78167	10328
2020-04-09	France	3881	541	82048	10869
2020-04-10	France	4286	1341	86334	12210
2020-04-11	France	4342	987	90676	13197

2020-04-12	France	3114	635	93790	13832
2020-04-13	France	1613	561	95403	14393
2020-04-14	France	2673	574	98076	14967
2020-04-15	France	5497	762	103573	15729
2020-04-16	France	2633	1438	106206	17167
2020-04-17	France	2641	753	108847	17920
2020-04-18	France	405	761	109252	18681



## 4 Vers un passage du pic ?

On s'intéresse ici à l'atteinte d'un pic du nombre de décès quotidiens liés à Covid-19. Là encore, l'axe temporel est recalé afin de permettre une comparaison entre pays :  $t = 0$  est positionné au moment où le nombre quotidien de morts dépasse 15.

```
threshold = 15
```

```
daily = {}
```

```
for country in countries:
```

```
    daily[country] = data_raw.query(f"location == '{country}' & new_deaths > {threshold}")
```

```
daily[countries[2]]
```

date	location	new_cases	new_deaths	total_cases	total_deaths
2020-03-14	France	785	18	3661	79
2020-03-16	France	924	36	5423	127

2020-03-17	France	1210	21	6633	148
2020-03-18	France	1097	27	7730	175
2020-03-19	France	1404	69	9134	244
2020-03-20	France	1861	128	10995	372
2020-03-21	France	1617	78	12612	450
2020-03-22	France	1847	112	14459	562
2020-03-23	France	1559	112	16018	674
2020-03-24	France	3838	186	19856	860
2020-03-25	France	2446	240	22302	1100
2020-03-26	France	2931	231	25233	1331
2020-03-27	France	3922	365	29155	1696
2020-03-28	France	3809	299	32964	1995
2020-03-29	France	4611	319	37575	2314
2020-03-30	France	2599	292	40174	2606
2020-03-31	France	4376	418	44550	3024
2020-04-01	France	7578	499	52128	3523
2020-04-02	France	4861	509	56989	4032
2020-04-03	France	2116	471	59105	4503
2020-04-04	France	5233	2004	64338	6507
2020-04-05	France	4267	1053	68605	7560
2020-04-06	France	1873	518	70478	8078
2020-04-07	France	3912	833	74390	8911
2020-04-08	France	3777	1417	78167	10328
2020-04-09	France	3881	541	82048	10869
2020-04-10	France	4286	1341	86334	12210
2020-04-11	France	4342	987	90676	13197
2020-04-12	France	3114	635	93790	13832
2020-04-13	France	1613	561	95403	14393
2020-04-14	France	2673	574	98076	14967
2020-04-15	France	5497	762	103573	15729
2020-04-16	France	2633	1438	106206	17167
2020-04-17	France	2641	753	108847	17920
2020-04-18	France	405	761	109252	18681

Le pic observé pour la France au 04/04 (jour 20 sur le graphique) correspond à la prise en compte instantanée de tous les décès en EHPAD, dont le décompte est disponible à partir du 01/02 (*cf.* données du gouvernement)

