



## **Etude statistique de la COVID-19 : analyse statistique de la situation de la pandémie et prévision économétrique des infectés au Sénégal**

Auteurs : A. DIAGNE<sup>1</sup> et O. A. KEBE<sup>2</sup>

### Résumé

Ce papier relate une analyse statistique de la situation de la pandémie de COVID-19 et une prévision économétrique des infectés. La pandémie a démarré à Diourbel en mars dernier et s'est poursuivie jusqu'à atteindre là-bas un pic vers la fin du mois de février. A la date du 02 juin 2020, Diourbel constitue 11% du total des cas positifs au Sénégal. Ensuite, Thiès a vu augmenter radicalement le nombre de cas infectés depuis le mois de mars, atteignant 5% du total de cas positifs au Sénégal. En même temps, et à des degrés moindres, la région de Dakar a petit à petit pris le relais. Dakar contient aujourd'hui plus de cas positifs que partout ailleurs au Sénégal (74%). Par contre, les régions du Sud à l'instar de Ziguinchor sont devenues l'épicentre de la pandémie à la fin du mois de mai avec 6% des cas positifs au Sénégal. Les régions de l'Est sont modestement touchées par ladite pandémie (environ 2% du total des cas positifs), de même que les autres régions du Centre (moins de 20 cas positifs seulement). Les estimations économétriques se sont effectuées en admettant que la pandémie de COVID-19 continue jusqu'au mois de décembre. Les résultats montrent que la durée de la pandémie ne dépasserait pas 300 jours au Sénégal, si durant les 10 mois, le nombre de nouveaux cas ne dépasse jamais 200, et en l'absence de vaccin et traitement efficace, avec un nombre de cas confirmés environ entre 10 000 et 55 000 personnes dont près de 1 000 morts au maximum au 31 décembre 2020.

**Mots clés : pandémie de COVID-19 ; étude statistique ; analyse statistique ; prévision économétrique ; Sénégal**

### Introduction

En décembre 2019, dans la province de Hubei et plus précisément dans la ville de Wuhan (en Chine centrale) est apparue la maladie à coronavirus, et ensuite officiellement nommée la COVID-19. Elle se propage ensuite partout dans le monde, le 13 janvier les Etats-Unis sont touchés par la maladie à coronavirus, le 24 janvier la France enregistre ses trois premiers cas. Le 30 janvier l'état d'urgence de santé publique de portée internationale est prononcé par l'OMS. La maladie continue de faire des ravages aussi bien en termes de personnes contaminées mais aussi en termes de décès par jour. Toutefois, le continent africain enregistre son premier cas le 15 février 2020, deux mois après l'apparition de la pandémie. Au Sénégal, la pandémie de COVID-19 débute officiellement le 02 mars 2020 avec un cas importé testé positif. On note une montée rapide du nombre de cas qui suit une tendance haussière, mais aussi le nombre de décès par jour continue sa progression haussière. Dès lors la question qui se pose est essentiellement de savoir à quel moment le pic de progression sera atteint ?

---

<sup>1</sup> M. Atoumane DIAGNE, Directeur adjoint chez Stat Agency

<sup>2</sup> M. Oumar A. KEBE, Responsable des enquêtes chez Stat Agency



Cette présente étude pose comme postulat principal que la pandémie de COVID-19 pourrait continuer jusqu'au mois de décembre. En utilisant aussi deux approches méthodologiques : l'analyse statistique de la pandémie et la prévision économétrique des infectés, les résultats montrent que la durée de la pandémie ne dépasserait pas 300 jours, si durant les 10 mois le nombre de nouveau cas ne dépasse jamais 200 et en l'absence de vaccin et traitement efficace, avec un nombre de cas confirmés entre 12 000 et 53 000 personnes dont 1 200 morts au maximum.

## 1 Méthodologie d'étude

Deux méthodes d'analyse seront utilisées dans cette étude : Analyse statistique de la pandémie et Prévision économétrique des infectés.

### 1.1 Analyse statistique

Les outils d'analyse de la Statistique descriptive (moyenne, médiane, écart type, minimum, maximum...) seront utilisés dans cette partie pour décrire la situation de 3 mois de pandémie. Des graphiques d'évolution de la pandémie s'ajouteront à ce panel de statistiques.

### 1.2 Prévision économétrique

Ces travaux se sont basés sur l'étude effectuée par Mathieu Grasland statisticien de l'ENSAI de Paris, en faisant une estimation paramétrique. L'idée générale est que les nouveaux cas infectés et les nouveaux morts du coronavirus durant une journée suivent la loi normale, car on remarque certaines propriétés similaires :

- Une montée rapide du nombre de cas
- Un pic caractéristique du centre temporel de l'épidémie.

L'objectif général est alors d'estimer la moyenne et l'écart type des deux courbes de loi normale. Sans équivoque, la moyenne de la loi normale va correspondre à un pic. Ce dernier va correspondre à un jour, qui sera la durée du premier pic de la pandémie. C'est ce pic qu'on veut calculer. Car si on multiplie par deux, on obtient directement la durée maximum de la pandémie au Sénégal.

Considérons  $x$  le numéro du jour,  $y$  le nombre de nouveaux cas et  $z$  le nombre de nouveaux morts durant la journée.

On compte à partir du 02 mars 2020 à chaque fois pour connaître la date correspondant au numéro du jour. Le pic de la première courbe correspond à la moyenne du nombre de nouveaux cas par jour  $\bar{y}$ . Actuellement, au moment du pic, c'est près de 180 nouveaux cas au Sénégal. Le pic de la seconde courbe correspond à la moyenne du nombre de nouveaux morts par jour  $\bar{z}$ .

En matière de prévisions, on pourra toujours essayer d'expliquer le paradoxe des résultats face à la réalité. Bien évidemment, on doit aussi toujours supposer le pire lorsqu'on fait des prévisions. Trois scénarii sont alors posés :

- Scénario tendanciel : nombre de nouveaux cas maîtrisés (en deçà de 200 cas)
- Scénario réaliste : beaucoup de nouveaux cas infectés en fin juin ou début juillet
- Scénario pessimiste : beaucoup de nouveaux cas infectés en fin août ou début septembre.

En effet, trois situations inattendues peuvent se produire, pouvant donner beaucoup plus de nouveaux cas (par exemple 500 nouveaux cas confirmés positifs par jour) :

- a) Les émigrés décident par tous les moyens de revenir dans leur pays natal en masse ;



- b) L'Etat décide brusquement de déconfiner les populations entraînant de nouvelles vagues ;
- c) L'Etat débute le dépistage massif soldé par de nouveaux cas positifs.

Une autre hypothèse dans cette méthodologie est que la pandémie perdure jusqu'au mois de décembre 2020 (donc 10 mois maximum de durée de pandémie).

## 2 Analyse statistique et Prévision économétrique

### 2.1 Analyse statistique de la situation de la COVID-19

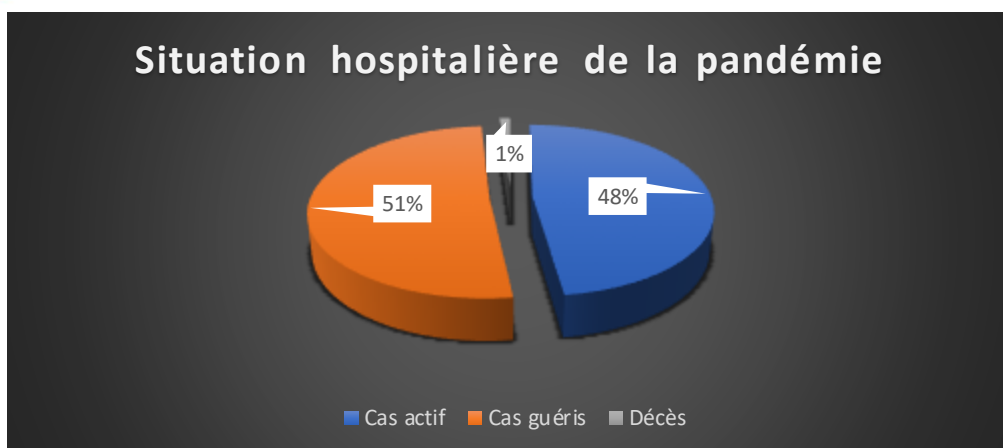
A la date du 02 juin 2020, la situation de la pandémie au Sénégal est la suivante :

**Tableau 1.** Statistiques descriptives de la situation du COVID-19

Statistiques	Nombre de nouveaux décès	Nombre de nouveaux cas
nbr.val	93	93
1st Qu.	0	7
3rd Qu.	1	75
min	0	0
max	4	177
range	4	177
sum	43	3836
median	0	15
mean	0,4623656	41,24731
SE.mean	0,08724008	4,288199
CI.mean.0.95	0,1732663	8,516734
var	0,7078074	1710,145
std.dev	0,8413129	41,35390
coef.var	1,819584	1,002584
skewness	2,013025	0,7559316
skew.2SE	4,025580	1,511686
kurtosis	3,847525	-0,4505965
kurt.2SE	3,885144	-0,4550022
normtest.W	0,6109148	0,8501280
normtest.p	2,483644e-14	2,900336e-08

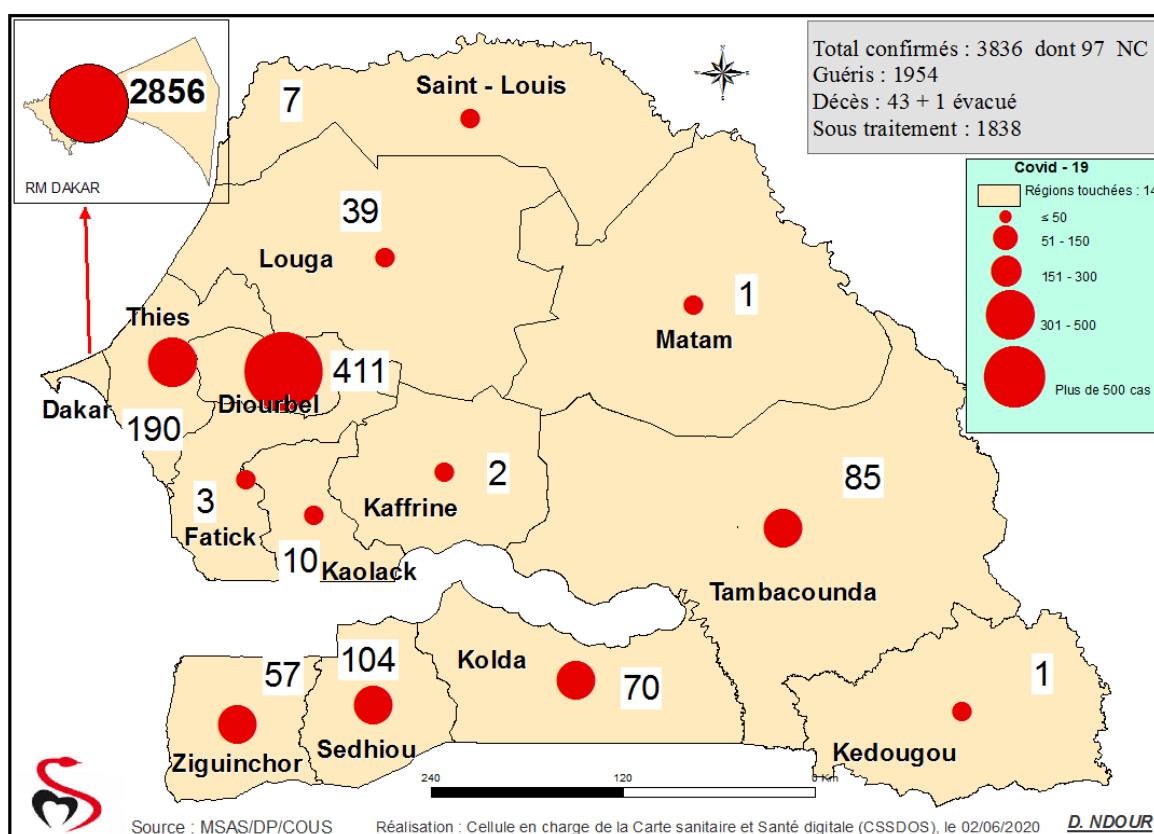
**Source :** Auteurs à partir des données de Our World In Data (Université d'Oxford), 02 juin 2020

Au total, à la date du 02 juin 2020, 43 décès ont été recensés au Sénégal avec 3836 cas confirmés du COVID-19. En moyenne, nous avons 41 cas confirmés et 0,46 décès par jour avec un écart type de 4,29 et 0,09 personnes respectivement. La médiane des nouveaux cas positifs journaliers correspond à 15 infectés tandis que pour les décès, elle est de l'ordre de 0,0001 morts. Cette situation se résume par un taux de létalité de 1% et un taux de guérison de 51%.



**Source** : Auteurs à partir des données de Our World In Data (Université d'Oxford), 02 juin 2020

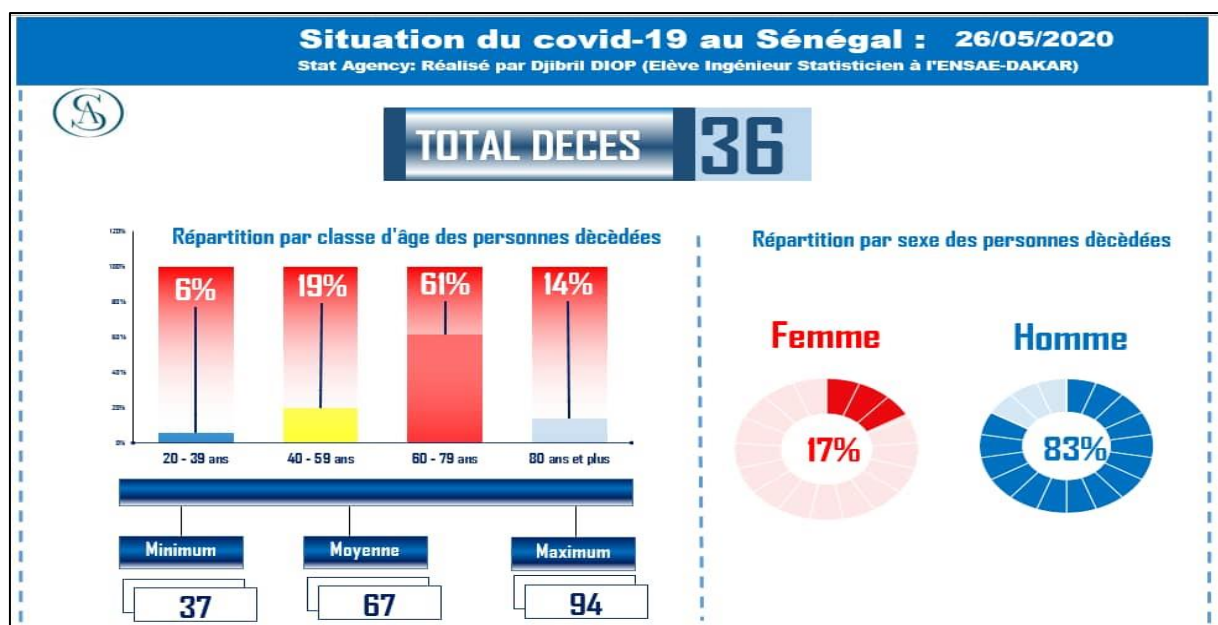
Une comparaison du total des cas positifs par région donne les valeurs consignées dans le graphique ci-contre. La pandémie a démarré à Diourbel en mars dernier et s'est poursuivie jusqu'à atteindre là-bas un pic local vers la fin du même mois. A la date du 02 juin 2020, Diourbel constitue 11% du total des cas positifs au Sénégal.



Ensuite, Thiès a vu augmenter radicalement le nombre de cas infectés depuis le mois de mars, atteignant 5% du total de cas positifs au Sénégal. En même temps, et à des degrés moindres, la région de Dakar a petit à petit pris le relais. Dakar contient aujourd'hui plus de cas positifs que partout ailleurs au Sénégal (74%). Par contre, les régions du Sud à l'instar de Ziguinchor sont devenues l'épicentre de la pandémie à la fin du mois de mai avec 6% des cas positifs au Sénégal. Les régions de l'Est sont



En faisant une analyse de la situation en termes de cas confirmés et de décès, à la fin du mois de mai, nous avons les graphiques suivants :







## Situation du covid-19 au Sénégal : 26/05/2020

Stat Agency; Réalisé par Djibril DIOP (Elève Ingénieur Statisticien à l'ENSAE-DAKAR)

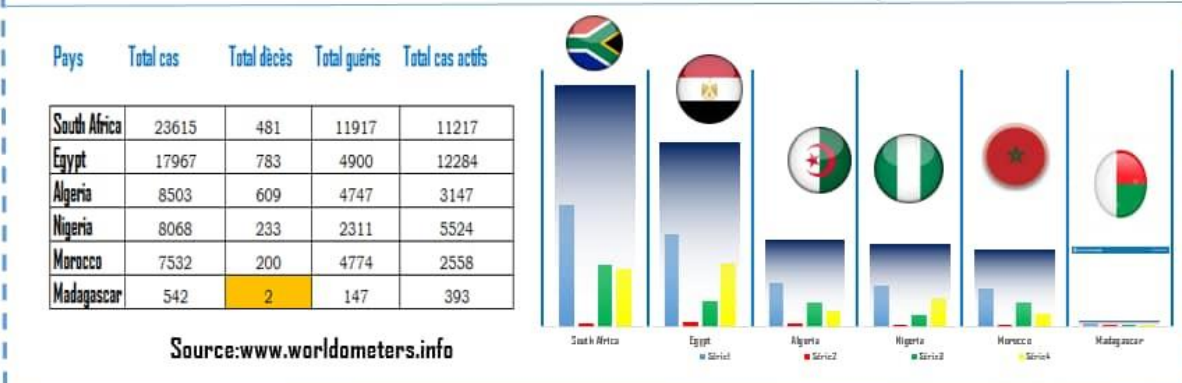
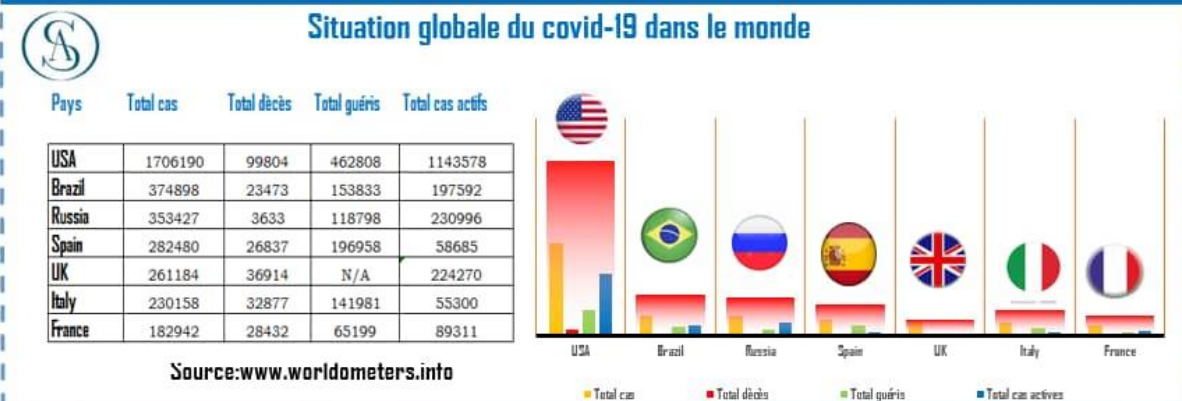


Source: Djibril DIOP, données du Ministère de la Santé et de l'Action Sociale (MSAS)

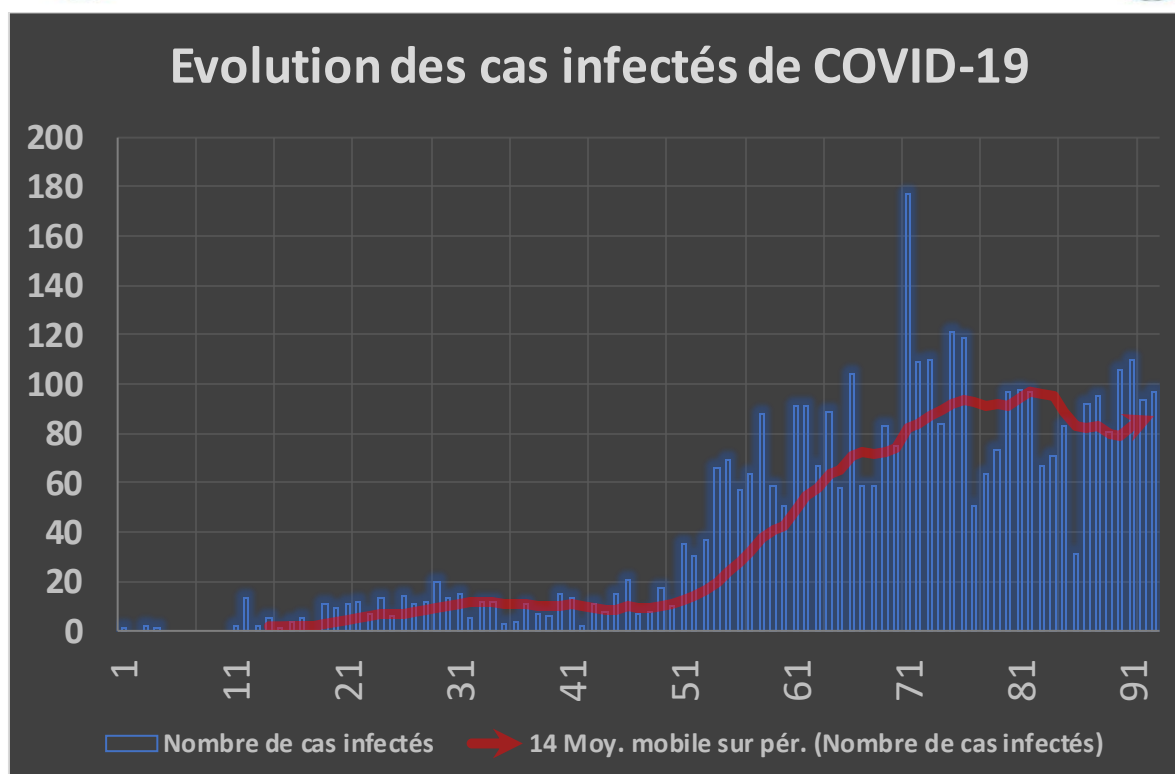
## Situation du covid-19 dans le monde :26/05/2020

Stat Agency; Réalisé par Djibril DIOP (Elève Ingénieur Statisticien l'ENSAE-DAKAR)

### Situation globale du covid-19 dans le monde



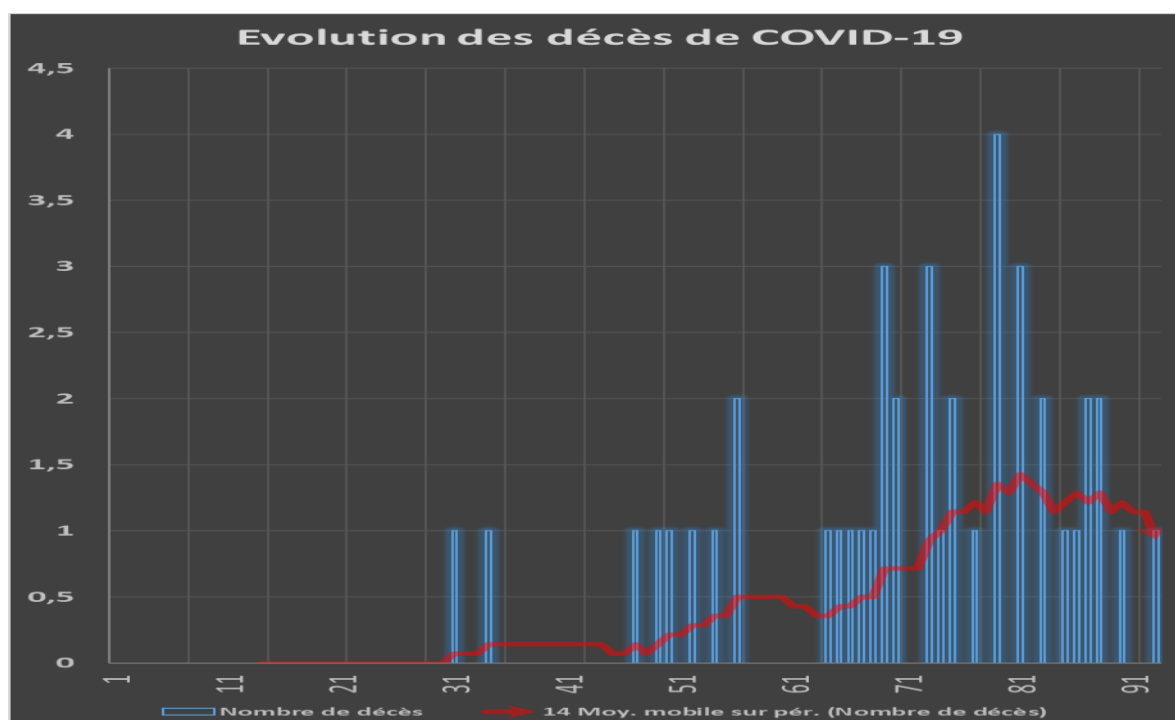
La tendance actuelle de la pandémie est à la stabilité du nombre de nouveaux cas comme le montre le graphique ci-après (environ 90 cas par jour). L'augmentation exponentielle des cas confirmés dès le mois de mars a continué jusqu'au mois d'avril. Au mois de mai, le nombre de nouveaux cas a une tendance stable comme l'illustre la courbe de moyenne mobile de période 14 jours (qui correspond à la durée théorique d'incubation de la maladie).



**Source** : Auteurs à partir des données de Our World In Data (Université d'Oxford), 02 juin 2020

\*En rouge Lissage par moyenne mobile de période 14

La situation des décès par jour est la même que celle des cas infectés comme on peut s'en apercevoir dans le graphique ci-bas (33 au mois de mai, 9 au mois d'avril, et zéro décès noté au mois de mars).



**Source** : Auteurs à partir des données de Our World In Data (Université d'Oxford), 02 juin 2020

\*En rouge Lissage par moyenne mobile de période 14



## 2.2 Prévision économétrique de la durée de la pandémie

Avec les statistiques descriptives du nombre de nouveaux cas et du nombre de nouveaux décès, le maximum de décès et d'infectés peut être estimé selon l'estimation du Total de Horvitz-Thompson (avec une répartition équiprobable des deux séries de données). Cette situation prend même en considération la possibilité d'une nouvelle vague de la pandémie<sup>3</sup> :

- $41,24731 \times 10 \times 30 = 12\,374$  cas confirmés en moyenne
- $0,4623656 \times 10 \times 30 = 139$  décès en moyenne
- $177 \times 10 \times 30 = 53\,100$  cas confirmés maximum
- $4 \times 10 \times 30 = 1\,200$  décès maximum

Les estimations paramétriques pour la courbe du nombre de nouveaux cas infectés (suivant une loi normale) donnent 150 pour la moyenne et 45 pour l'écart type, avec 0.05512 d'erreur d'estimation (voir Annexe). Les estimations pour la courbe du nombre de nouveaux morts (suivant une loi normale) donnent 150 pour la moyenne et 53,3 pour l'écart type, avec 0.1636622 d'erreur d'estimation. Avec cela, on peut calculer le minimum et le maximum du nombre de cas infectés et décès prévus pour chaque date (de 0 à 300 jours).

Les estimations se sont effectuées en admettant que la situation de la COVID-19 continue jusqu'au mois de décembre (300 jours). Les résultats montrent qu'on n'aurait plus de nouveaux cas infectés au-delà de cette période (voir Annexe). On peut conclure que la durée du COVID-19 ne dépasserait pas 300 jours au Sénégal, si durant les 10 mois, le nombre de nouveaux cas ne dépasse jamais 200, et en l'absence de vaccin et traitement efficace contre la COVID-19 en 2020. Sinon on l'aurait remarqué durant les 10 premiers mois, et on poserait un nouveau scénario imprévisible dans ce cas. Le pic sera atteint vers la fin du mois de juin et le début du mois de juillet dans le scénario tendanciel comme on peut s'en apercevoir dans les graphes des estimations.

La durée de la pandémie pourrait donc atteindre 300 jours avec un nombre de cas confirmés entre 12 000 et 53 000 personnes dont près de mille morts au maximum. Après ces 10 mois, on aurait des cas confirmés encore qui seront négligeables quand même (en 2021). Au-delà de cette période de 300 jours, et en l'absence de vaccin et traitement efficace contre la COVID-19, il y aurait encore des gens en hospitalisation ou des nouveaux cas confirmés d'une à 50 personnes environ par jour. Le nombre de nouveaux infectés, ce ne sera pas zéro cas, tout de suite. Pour la Chine par exemple, après le confinement total, il y a eu un petit nombre de cas des jours et des jours.

## Conclusion

Au mois de juin, la bonne nouvelle c'est que le nombre de nouveaux cas est maîtrisé (avec un taux de positivité des tests de moins 20% les deux derniers mois).

---

<sup>3</sup> C'est la moyenne et le maximum des nouveaux cas et nouveaux décès à la date du 02 juin qui sont prises dans ces calculs. On devrait actualiser chaque jour ces données pour que les estimateurs convergent vers les vraies valeurs.





Quant aux décideurs publics, ils sont sur la bonne voie et ils doivent continuer ces mesures barrières jusqu'à mi-juillet. Si l'État continue ces mêmes mesures barrières jusqu'à mi-août, on peut dire qu'il n'y aurait pratiquement plus de nouveaux cas au Sénégal au-delà de décembre 2020.

Les recommandations préventives suivantes restent en vigueur pour le deuxième semestre de 2020 (juillet-décembre) :

1. Mise en quarantaine de tout immigrant mettant ses pieds sur le territoire national
2. Renforcement de la sécurité au niveau des frontières terrestres
3. Désinfection de tous les lieux suspects de contamination
4. Un plan de mise en quarantaine des émigrés de retour dans leur pays et accompagnement économique et social
5. Continuation des mesures barrières et du couvre-feu
6. Reprise progressive des investissements publics dans la mesure où ils permettent d'avoir des emplois et de la croissance économique dans les pays en développement
7. Relance des activités pour les produits de forte valeur ajoutée et de grande nécessité.

Les émigrés surtout non régularisés restent dans une situation incertaine, surtout que les pays développés sont en récession.

Enfin, à partir du mois de juillet, le Sénégal devrait agir comme la Chine aux mois de mars-juin, c'est-à-dire ce qu'ils font actuellement.

La prochaine question à laquelle il faudra penser est, quand introduire la fin du couvre-feu sans redouter de relancer l'épidémie ? Et les politiques à mettre en place pour une reprise économique efficace ?

Dans les prochaines publications, nous pouvons tenter différentes modélisations avec de nouvelles lois, toujours dans le souci de donner au mieux la meilleure des estimations.

#### Références

Mathieu Grasland, Estimation de la durée de l'épidémie du COVID-19, Document de travail, ENSAI de Paris. 2020.

Klaus Nordhausen and Hannu Oja. Robust Nonparametric Inference. Annual Reviews of Statistics and its Application. 2018; 5: 473-500.

#### Sites internet :

[https://www.lemonde.fr/planete/live/2020/02/26/le-bilan-du-coronavirus-en-direct-undeuxieme-patient-est-mort-a-paris\\_6030908\\_3244.html](https://www.lemonde.fr/planete/live/2020/02/26/le-bilan-du-coronavirus-en-direct-undeuxieme-patient-est-mort-a-paris_6030908_3244.html) consulté le 26 février 2020

<https://www.worldmeters.info/coronavirus/> consulté le 27 mai 2020

<https://www.banquemonddiale.org/fr/news/statement/2020/03/27/world-bank-group-president-david-malpass-remarks-to-the-international-monetary-and-financial-committee> consulté le 27 mai 2020

<https://www.ourworldindata.org/coronavirus-source-data> consulté le 02 juin 2020

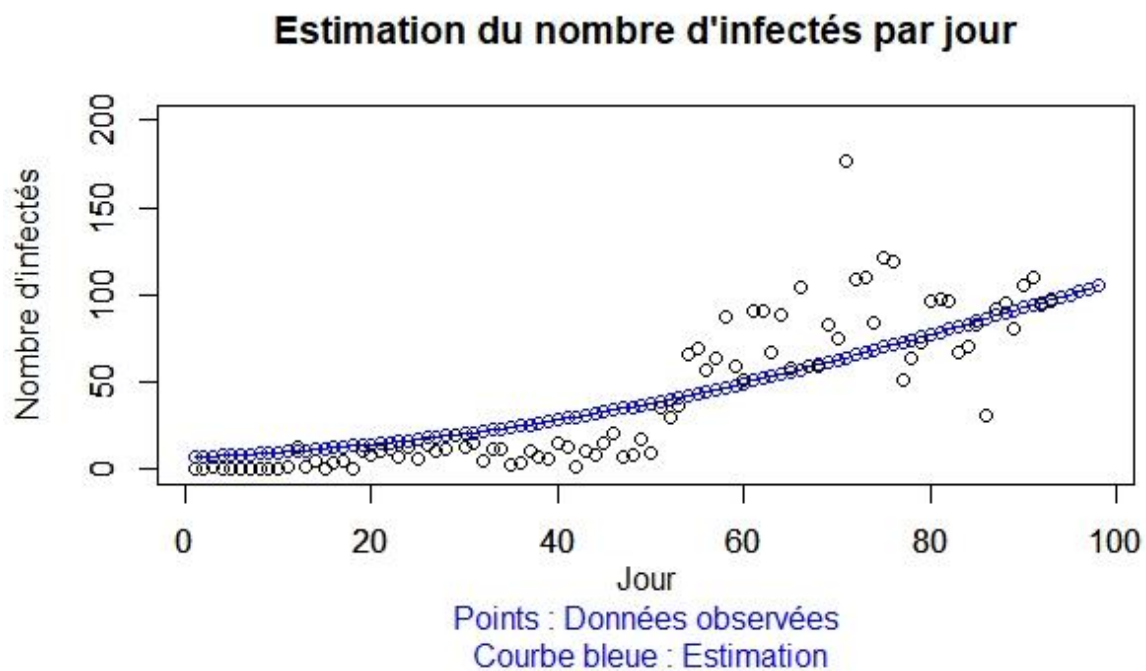
<https://www.sante.gouv.sn> consulté le 02 juin 2020



## Prévision des infectés :

### Scénario 1 pic tendanciel

Le pic de contaminations dues à l'épidémie serait le 150ème jour coïncidant approximativement à la moitié de la durée d'estimation qui est de 10 mois (cf. Méthodologie). L'estimation du nombre total de cas est de 26475 avec un taux d'erreur de 0.05512.



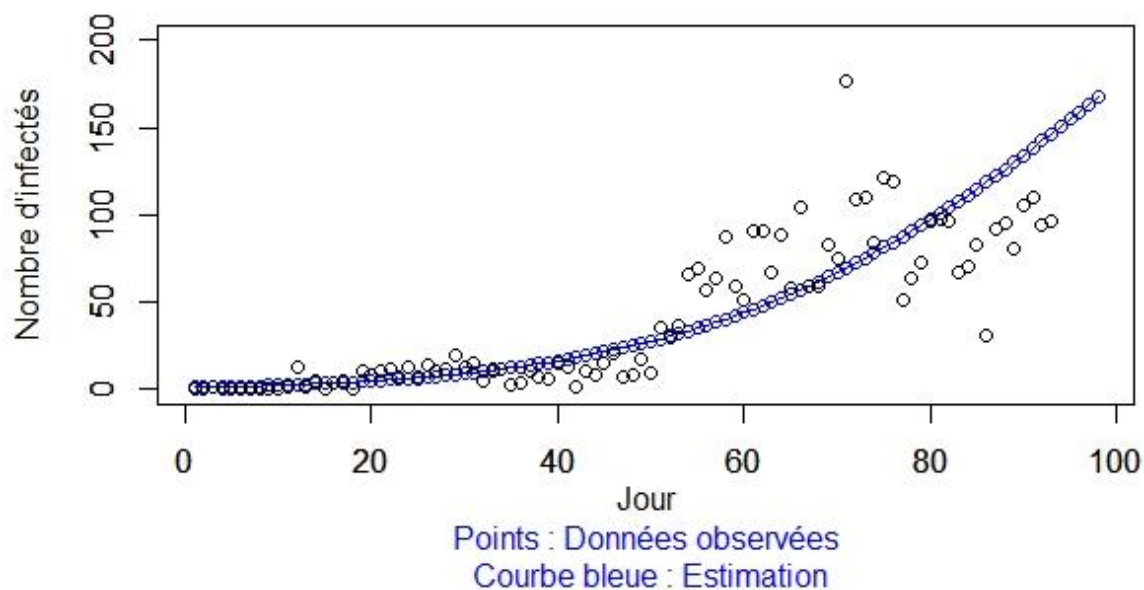
**Source :** Auteurs à partir des données de Our World In Data (Université d'Oxford), 02 juin 2020

### Scénario 2 pic fin juin juillet

Le pic de contaminations dues à l'épidémie serait le 150ème jour coïncidant approximativement entre la fin du mois de juin et le début du mois de juillet. L'estimation du nombre total de cas est de 36818 avec un taux d'erreur de 0.06302188.



### Estimation du nombre d'infectés par jour

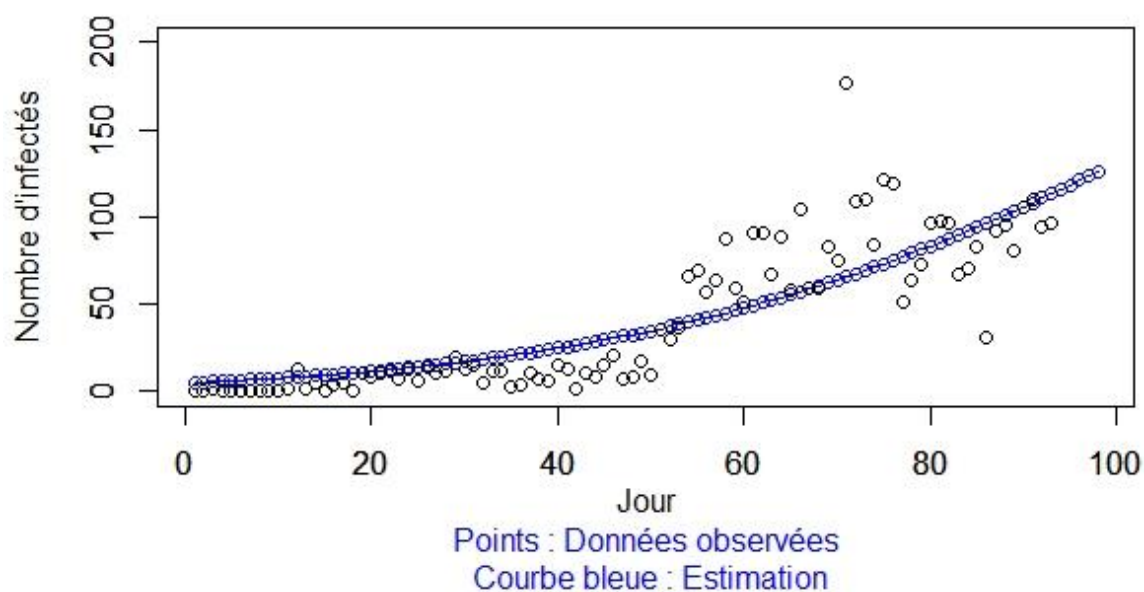


**Source** : Auteurs à partir des données de Our World In Data (Université d'Oxford), 02 juin 2020

#### Scénario 3 pic août septembre

Le pic de contaminations dues à l'épidémie serait le 180ème jour coïncidant approximativement entre la fin du mois d'août et le début du mois de septembre. L'estimation du nombre total de cas est de 46396 avec un taux d'erreur de 0.0561017.

### Estimation du nombre d'infectés par jour



**Source** : Auteurs à partir des données de Our World In Data (Université d'Oxford), 02 juin 2020

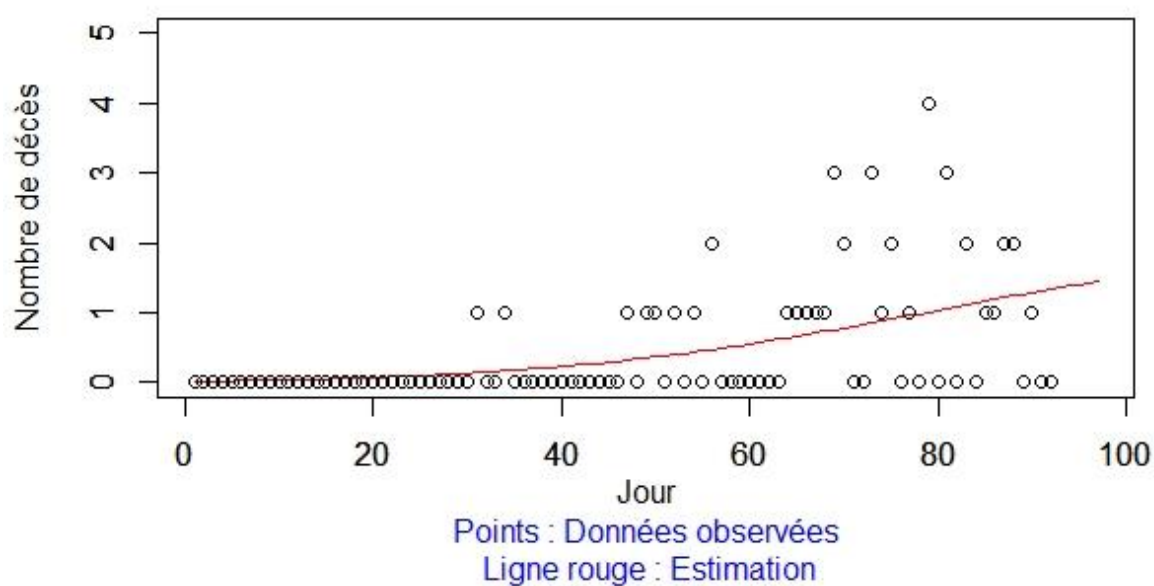


## Prévision des décès :

### Scénario 1 pic tendanciel

Le pic de décès dus à l'épidémie serait le 150ème jour coïncidant à la moitié de la durée d'estimation qui est de 10 mois (cf. Méthodologie). L'estimation du nombre total de morts est de 303 avec un taux d'erreur de 0.1636622.

### Estimation du nombre de décès par jour



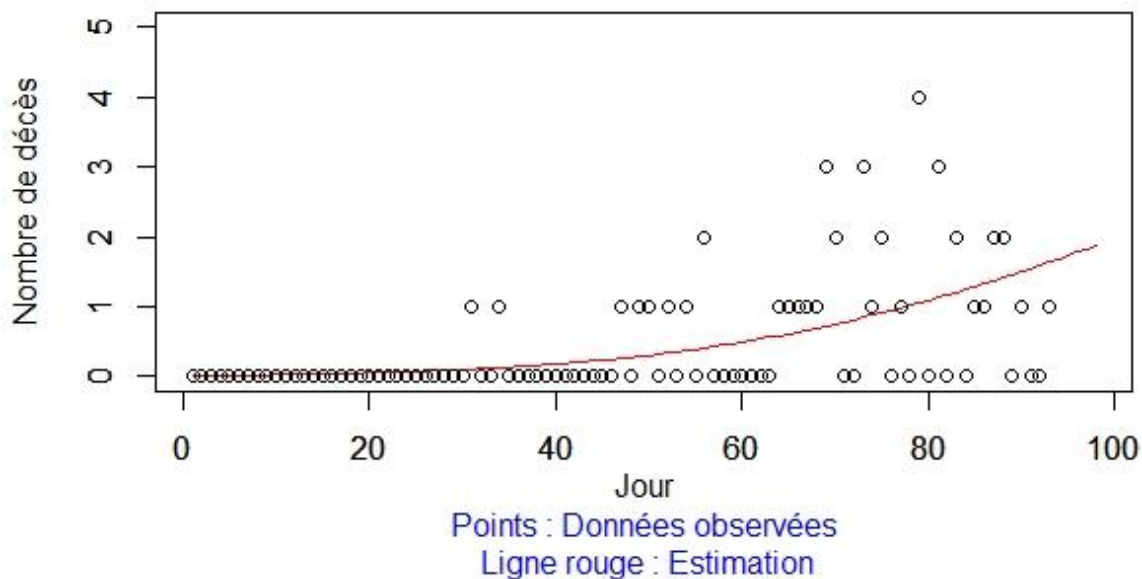
**Source :** Auteurs à partir des données de Our World In Data (Université d'Oxford), 02 juin 2020

### Scénario décès pic fin juin juillet

Le pic de décès dus à l'épidémie serait le 150ème jour coïncidant approximativement entre la fin du mois de juin et le début du mois de juillet. L'estimation du nombre total de morts est de 413 avec un taux d'erreur de 0.1664001.



## Estimation du nombre de décès par jour

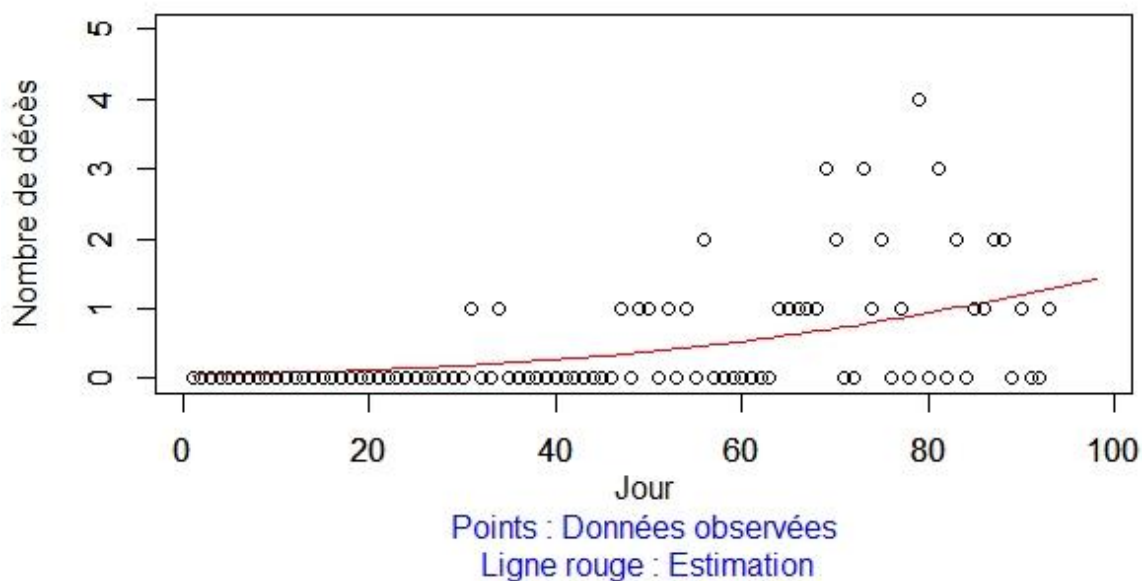


**Source** : Auteurs à partir des données de Our World In Data (Université d'Oxford), 02 juin 2020

### Scénario décès pic août septembre

Le pic de décès dus à l'épidémie serait le 180ème jour coïncidant approximativement entre la fin du mois d'août et le début du mois de septembre. L'estimation du nombre total de morts est de 534 avec un taux d'erreur de 0.1641996.

## Estimation du nombre de décès par jour



**Source** : Auteurs à partir des données de Our World In Data (Université d'Oxford), 02 juin 2020