

COVID-19 EN FRANCE

SUIVI ET ANALYSE



Arthur Ba

DECEMBRE 2020

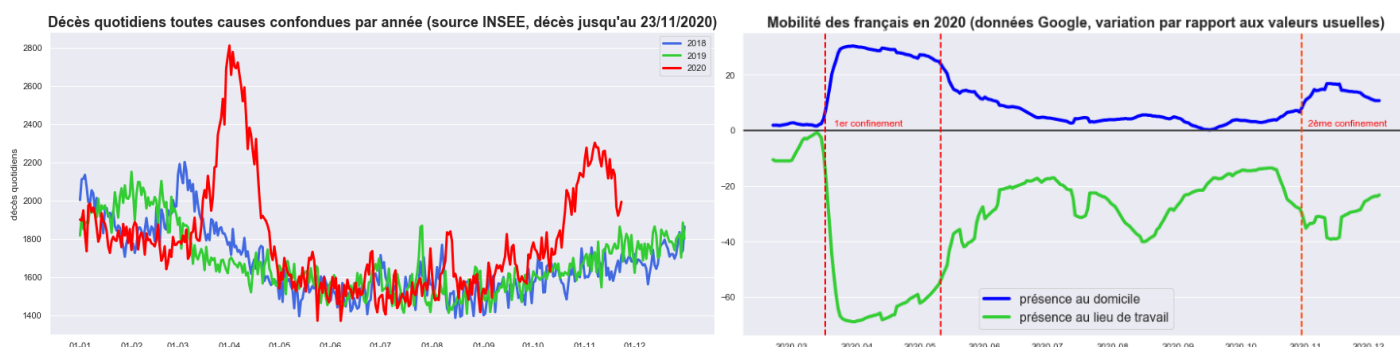


INTRODUCTION

Une année 2020 marquée par la pandémie

L'épidémie de coronavirus, ou COVID-19, a changé le quotidien de tous, d'une manière ou d'une autre. La violence de la première vague a mis le monde à l'arrêt, et aujourd'hui encore, notre économie et notre mode de vie sont bouleversés, bien que le pire de la deuxième vague soit derrière nous.

En France, on recense près de 60,000 décès attribués au COVID-19 à la mi-décembre. L'INSEE enregistre cette année une nette surmortalité, nettement visible durant le pic des deux vagues épidémiques. Près de 2,5 millions de personnes ont été testés positives, mais ce chiffre est probablement très éloigné de la réalité, puisque le nombre de cas positifs asymptomatiques non testés n'est pas mesurable.



Si la découverte d'un vaccin efficace en cette fin d'année offre des perspectives meilleures pour 2021, il est certain que nous devrons encore vivre avec le virus pendant au moins plusieurs mois.

Enjeux pour les semaines à venir

Dans ce rapport, nous faisons un point sur l'évolution de l'épidémie en France, et apportons des éléments permettant de juger sur les suites possibles de l'épidémie à court-terme. Nous essayons également de déterminer quels sont les facteurs expliquant pourquoi certains départements français sont plus touchés que d'autres par le virus.

SOMMAIRE

A – BILAN AU NIVEAU NATIONAL 4

- 1 – Lente décrue des hospitalisations, avant un rebond ?4
- 2 – Des différences marquées selon l'âge et le sexe5
- 3 – Les nouveaux cas à nouveau en hausse6

B – TENDANCES 7

- 1 – Les objectifs du 15/12 seulement en partie remplis..... 7
- 2 – Deux vagues aux caractéristiques différentes 7
- 3 – Une dégradation de la situation est à craindre8
- 4 – Projections sur 2 semaines.....9

C –ANALYSE PAR DEPARTEMENT 10

- 1 – Taux d'incidence par département 10
- 2 – Disparités entre les deux vagues selon les départements 11
- 3 – Les hospitalisations hebdomadaires comme indicateur 12

D – ETUDE DE CORRELATIONS 13

- 1 – Variance de la température 13
- 2 – Densité de la population 14
- 3 – Taux d'urbanisation 14
- 4 – Régression linéaire multiple 15

CONCLUSION 16

A - BILAN AU NIVEAU NATIONAL

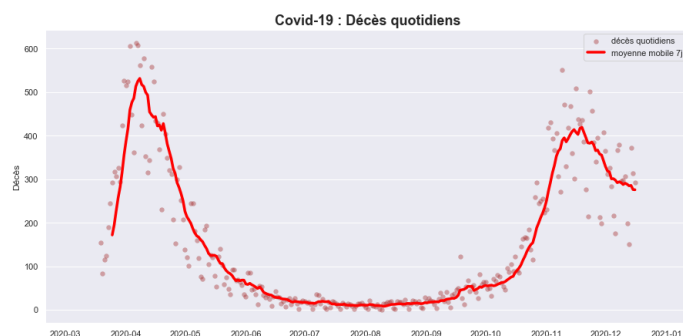
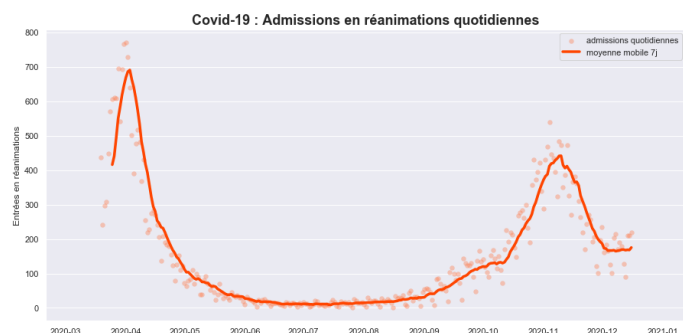
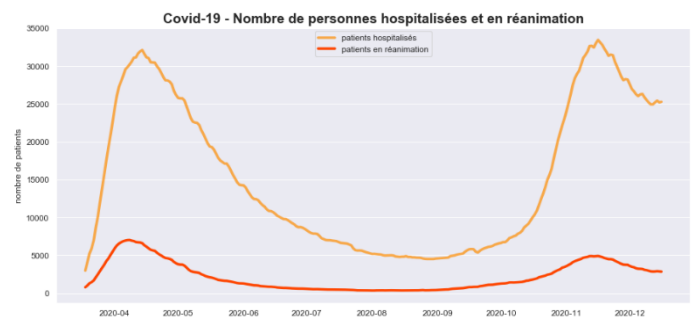
1 – Lente décrue des hospitalisations, avant un rebond ?

Après avoir atteint un maximum similaire à celui de la première vague, le nombre de personnes hospitalisées en France a constamment diminué jusqu'à début décembre. Néanmoins ce chiffre plafonne désormais autour de 25,000 personnes. Le nombre d'individus actuellement en soins intensifs suit la même tendance avec 2900 patients en réanimation à la mi-décembre.

L'analyse des flux entrants, notamment les nouvelles admissions quotidiennes montre que le nombre de patients admis ne diminue plus, et stagne à 180 entrées par jour. On observe même un sursaut des entrées depuis le 16 décembre, qui est pour le moment compensé par des flux sortants importants

Pour ce qui est des décès en milieu hospitalier, la décrue de la seconde vague est pour le moment constante, avec un nombre de décès quotidiens se situant en moyenne à 280 morts par jour. Le pic de la seconde vague a été atteint le 9 novembre avec 551 décès recensés.

Il est à craindre que cette diminution soit freinée dans les prochains jours. En effet, on observe un décalage d'environ deux semaines entre une évolution des hospitalisations et un impact similaire sur les décès. Il faut donc s'attendre à une stabilisation des décès quotidiens durant les 2 dernières semaines de l'année.

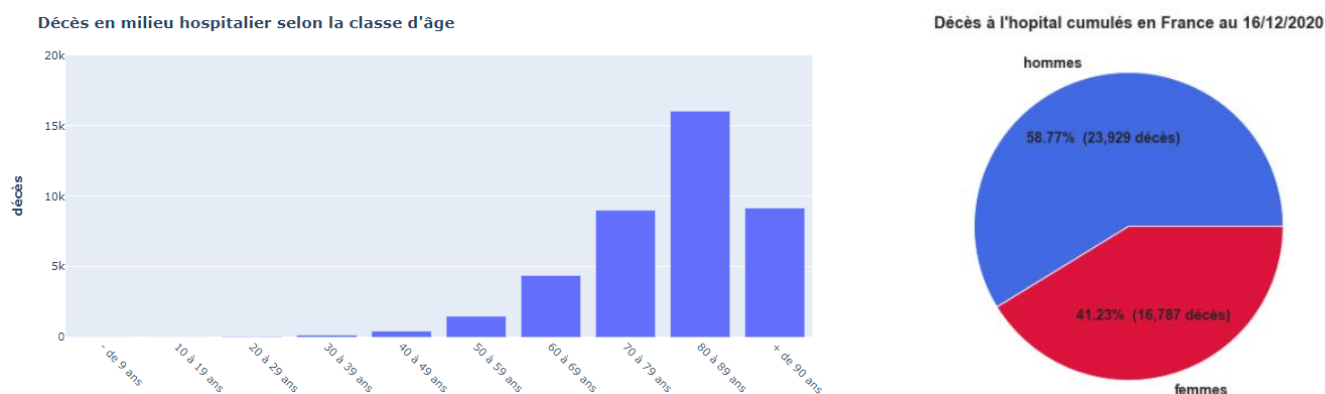


SOURCES :

Les données présentées ici proviennent des fichiers open data créés par Santé Publique France (SPF) et accessibles publiquement sur le site data.gouv.fr

A - BILAN AU NIVEAU NATIONAL

2 – Des différences marquées selon l'âge et le sexe



Les cas graves de COVID-19 touchent en grande majorité les personnes âgées, tandis que sa létalité est extrêmement faible. 94,8% des décès en milieu hospitalier concerne des patients âgés de plus de 60 ans. La catégorie des 80 à 89 ans est la plus touchée avec plus de 16,000 décès à la mi-décembre. Les décès de patients de moins de 40 ans constitue moins de 1% des décès enregistrés. Le virus est d'avantage létal pour les hommes que pour les femmes. Près de 6 patients décédés sur 10 sont des hommes. Cette différence peut s'expliquer par le fait que les individus masculins ont tendance à présenter des facteurs de comorbidité plus souvent que les femmes.



La carte de chaleur du taux d'incidence par classe d'âge, semaine par semaine, met en évidence la propagation du virus lors de la seconde vague. Les 20-29 ans ont été les premiers touchés par une hausse des cas positifs, durant l'été 2020. Ils ont par la suite contaminé les autres classes d'âge. Au plus fort de la seconde vague, l'incidence a dépassé les 800 cas pour 100,000 personnes chez les plus de 90 ans, et était supérieur à 400 pour toutes les classes d'âge hors moins de 9 ans. Les taux d'incidence ont, à la mi-décembre, retrouvé des valeurs antérieures au début de la seconde vague, même si l'incidence chez les plus de 90 ans reste élevée. Ce dernier chiffre est à pondérer par le fait que les personnes très âgées, et donc à risque se font relativement plus testées que les autres classes d'âge, ce qui explique en partie une incidence plus élevée.

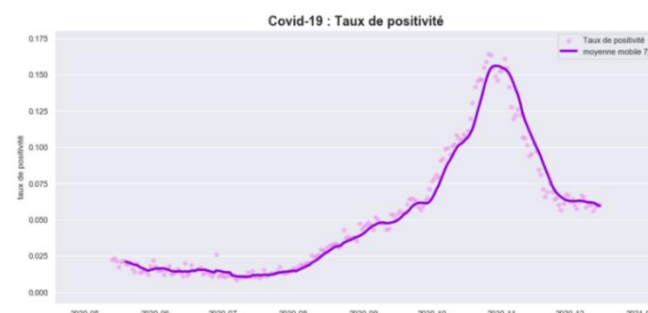
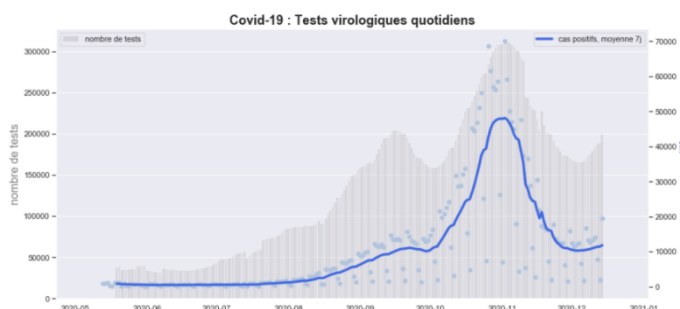
A - BILAN AU NIVEAU NATIONAL

3 – Les nouveaux cas à nouveau en hausse

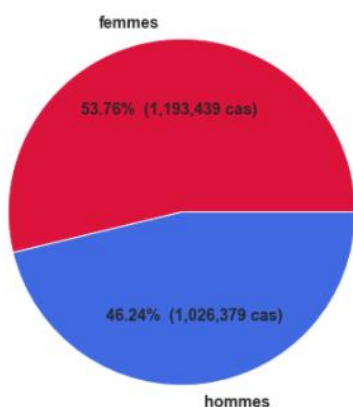
Après avoir atteint le pic de la seconde vague (70,063 cas positifs le 1^{er} novembre), les nouveaux cas confirmés ont diminué progressivement durant le mois de novembre. Néanmoins un plateau semble avoir été atteint autour de 10,000 nouveaux cas par jour, et à la mi-novembre la tendance est de nouveau à la hausse (11,807 cas en moyenne glissante 7j au 17 décembre).

Il est important de prendre de compte le nombre de tests réalisés sur le territoire, qui lui aussi est en hausse sur le mois de décembre, et contribue donc en partie à la hausse du nombre de cas.

Le taux de positivité (nombre de cas rapporté au nombre de tests) est lui aussi resté stable depuis le début du mois, mais ne montre pas de signe de rebond. Cet indicateur tend même à diminuer à la mi-décembre, en repassant sous la barre des 6% (en moyenne glissante)



Cumul des cas confirmés au 17/12/2020



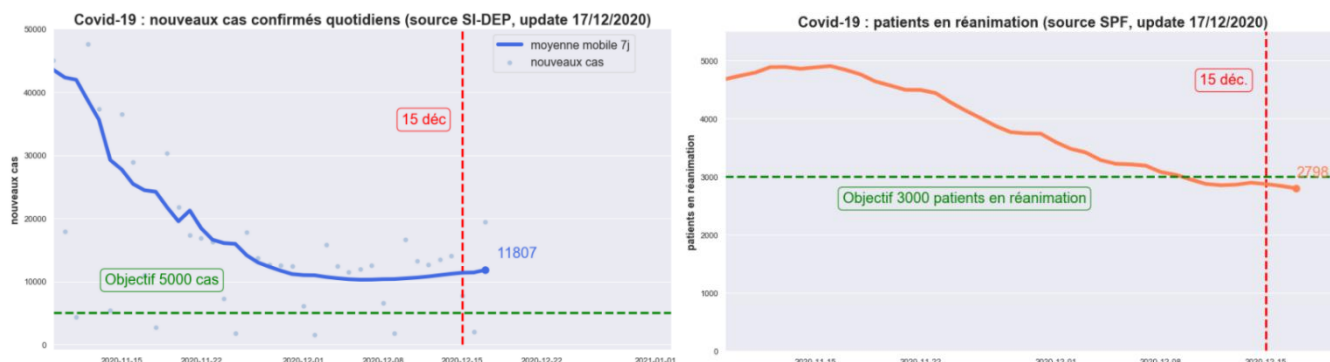
POURQUOI LE NOMBRE DE NOUVEAUX CAS VARIE SELON LES SOURCES ?

Il faut faire la distinction entre les nouveaux cas annoncés par date de résultat (jour où le test est confirmé positif) et date de prélèvement (jour où le test a été effectué).

Nous affichons ici les données de test par date de prélèvement

B – TENDANCES

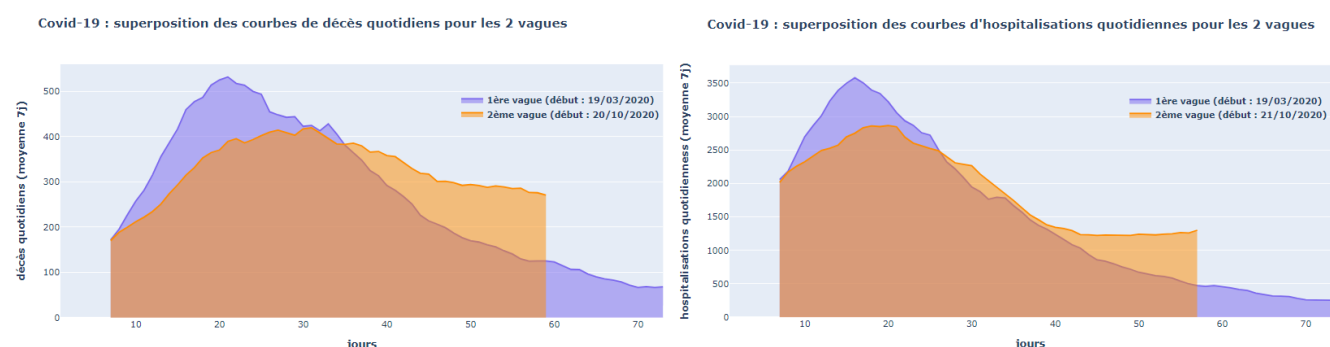
1 – Les objectifs du 15/12 seulement en partie remplis



L'objectif formulé fin novembre par le gouvernement français était de passer sous le seuil des 5000 nouveaux cas par jour avant le 15 décembre. Cet objectif n'a pas été rempli, et comme dis plus haut, le nombre de cas quotidiens dépasser les 11800 par jours en moyenne glissante à la mi-décembre, et augmente ces derniers jours.

En revanche l'objectif de 3000 patients maximum en réanimation a été réalisé, puisque le territoire comptait moins de 2900 patients en soins intensifs au soir du 15 décembre.

2 – Deux vagues aux caractéristiques différentes

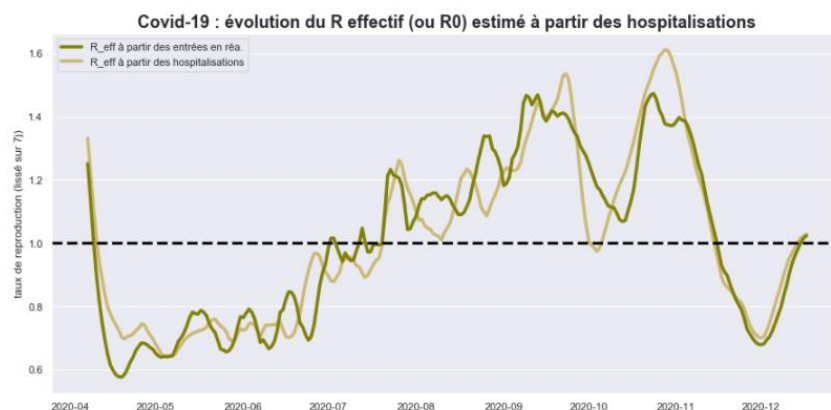


La seconde vague de l'épidémie a été moins violente que la première. Aussi bien pour les hospitalisations que pour les décès, le pic de début novembre est 20% inférieur à celui d'avril. En revanche, la seconde vague met bien plus de temps à ralentir. La stagnation des chiffres depuis début décembre n'a pas été observée durant la première vague, où la baisse a été continue.

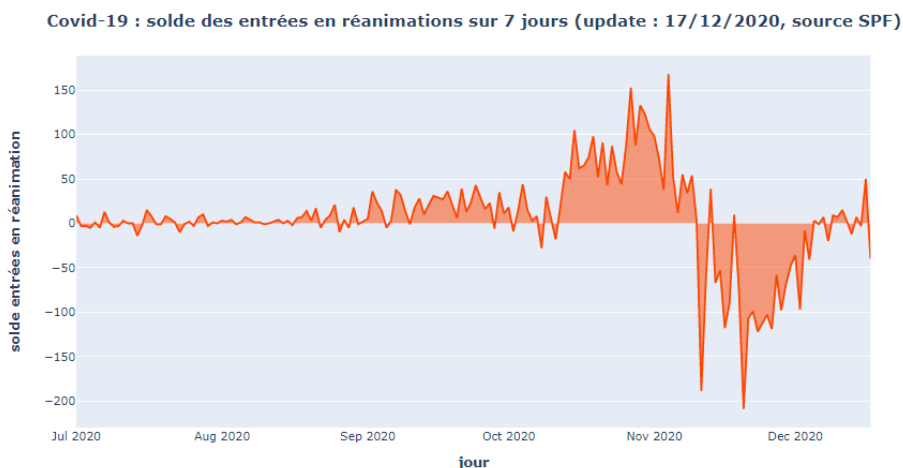
Cette superposition met en évidence le phénomène d'aplatissement de la courbe, qui est le résultat de gestes barrières renforcés en fin d'année ainsi que la mise en place de mesures restrictives plus en amont de la recrudescence des hospitalisations.

B – TENDANCES

3 – Une dégradation de la situation est à craindre



Le taux de reproduction du virus (R effectif, ou encore R0) a selon nos estimations atteint un pic à 1,6 (calculé à partir des hospitalisations) fin octobre. Il a diminué tout au long du mois de novembre pour s'établir autour de 0,7. Il est cependant en hausse constante depuis début décembre, et est de nouveau supérieur à 1 depuis la mi-décembre. Cela signifie que nous entrons potentiellement dans une nouvelle phase de propagation du virus dans la population.



Un des indicateurs les plus intéressants à suivre pour repérer un rebond éventuel de l'épidémie est le solde des entrées en réanimation par rapport au même jour de la semaine précédente. Ce solde était négatif et en baisse durant le mois de novembre, reflétant la phase descendante de la seconde vague. L'indicateur s'est néanmoins rapproché de zéro en décembre, et on observe à nouveau à la mi-décembre des valeurs positives, soit une tendance à la hausse des entrées en réanimations.

Si le déconfinement du 15 décembre et les fêtes contribueront à une nouvelle poussée du virus, ce solde nous permettra détecter le rebond rapidement, notamment si les valeurs dépassent le seuil de 100 entrées en réanimation supplémentaires par rapport à sept jours auparavant.

B – TENDANCES

4 – Projections sur 2 semaines

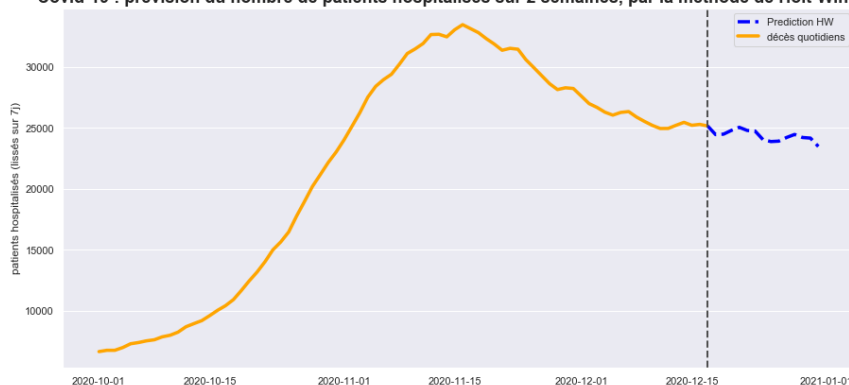
A l'aide d'un lissage exponentielle double (méthode de Holt-Winters), nous avons réalisé des prédictions sur un horizon de 14 jours, afin d'estimer les tendances pour différentes données hospitalières, présentées ci-dessous. Ces projections sont basées sur les données passées, et peuvent être très éloignées de la réalité. Néanmoins, elles nous permettent d'avoir une idée du rebond ou non de l'épidémie. Si au fil des jours les chiffres observés divergent nettement par rapport à la projection, on pourra conclure à un changement de dynamique dans la circulation du virus.

Nous projetons pour les prochains jours une légère diminution du nombre de personnes hospitalisées pour COVID-19. Cette baisse devrait être très lente, avec un rythme similaire à celui observé depuis début décembre.

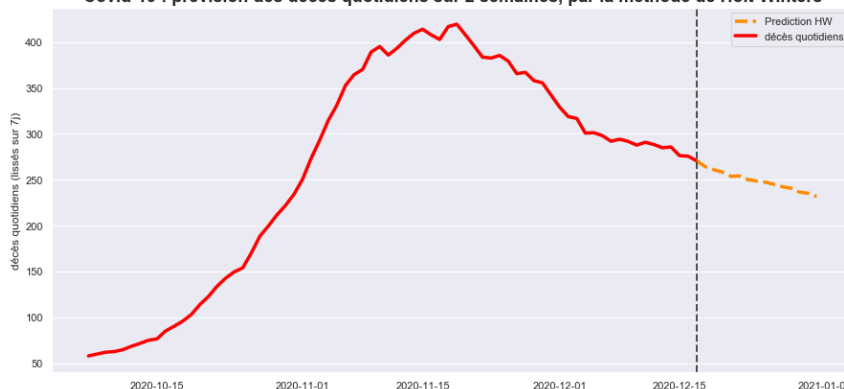
Les admissions en réanimation devraient rester stable selon la tendance actuelle. Les flux entrants en soins intensifs devraient ainsi rester proche de 170 admissions par jour.

Les décès quotidiens devraient quant à eux poursuivre leur décrue. Selon cette projection, le pays comptera moins de 230 décès par jour en moyenne glissante 7j d'ici la fin de l'année.

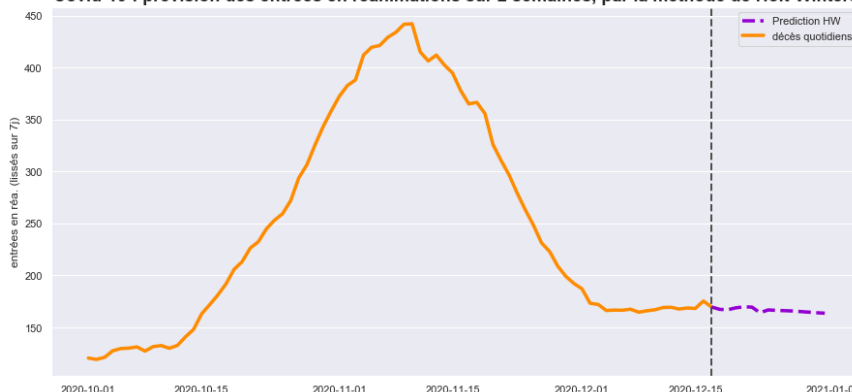
Covid-19 : prévision du nombre de patients hospitalisés sur 2 semaines, par la méthode de Holt-Winters



Covid-19 : prévision des décès quotidiens sur 2 semaines, par la méthode de Holt-Winters



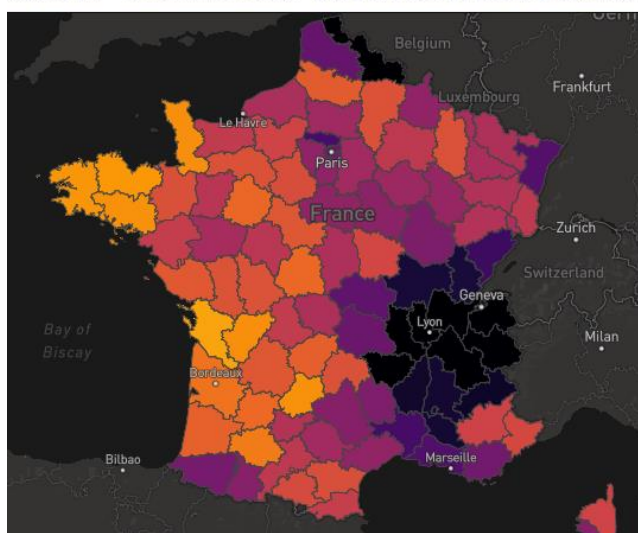
Covid-19 : prévision des entrées en réanimations sur 2 semaines, par la méthode de Holt-Winters



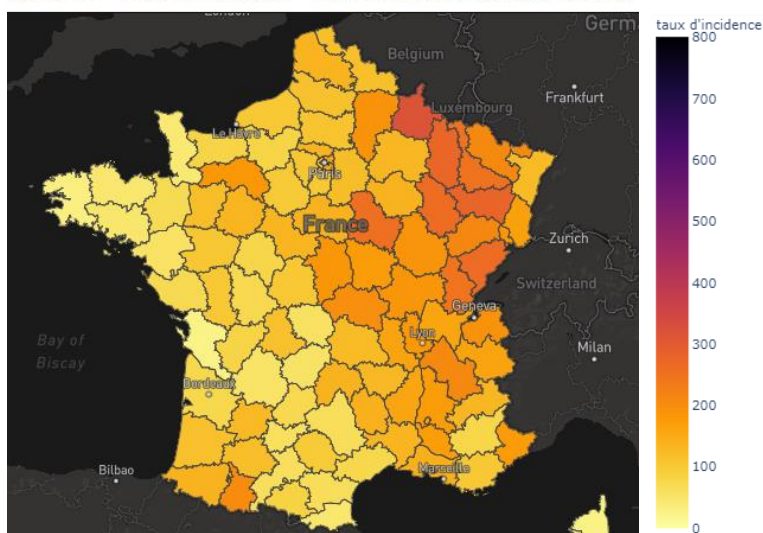
C – ANALYSE PAR DÉPARTEMENT

1 – Taux d'incidence par département

Covid-19 - Taux d'incidence - 01/11/2020 (7 jours glissants)



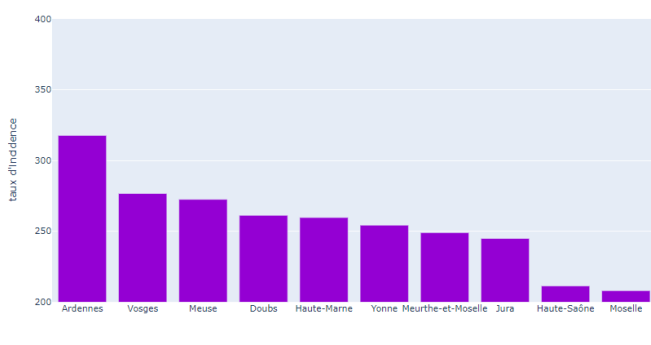
Covid-19 - Taux d'incidence - 14/12/2020 (7 jours glissants)



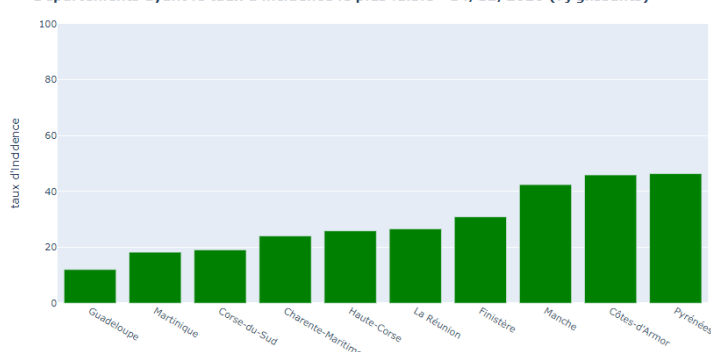
Il y a depuis le début de la pandémie une forte disparité entre départements. Au plus fort de la seconde vague, les départements du Sud Est, notamment ceux de la région Auvergne-Rhône-Alpes ont connu les taux d'incidence les plus élevés. La Haute-Savoie a notamment dépassé un taux d'incidence de 1100 en moyenne glissante. D'autres régions ont été relativement épargnées, comme la Bretagne. Les départements bretons n'ont pas dépassé 200 cas pour 100,000 habitants.

A la mi-décembre l'incidence a fortement diminué dans tous les départements français. Ce taux reste néanmoins partout bien plus élevé que les valeurs observées cette été. On constate depuis quelques jours une nouvelle hausse dans certains départements, notamment dans les Ardennes, et dans une moindre mesure les autres départements du Grand-Est.

Départements ayant le taux d'incidence le plus élevé - 14/12/2020 (7j glissants)

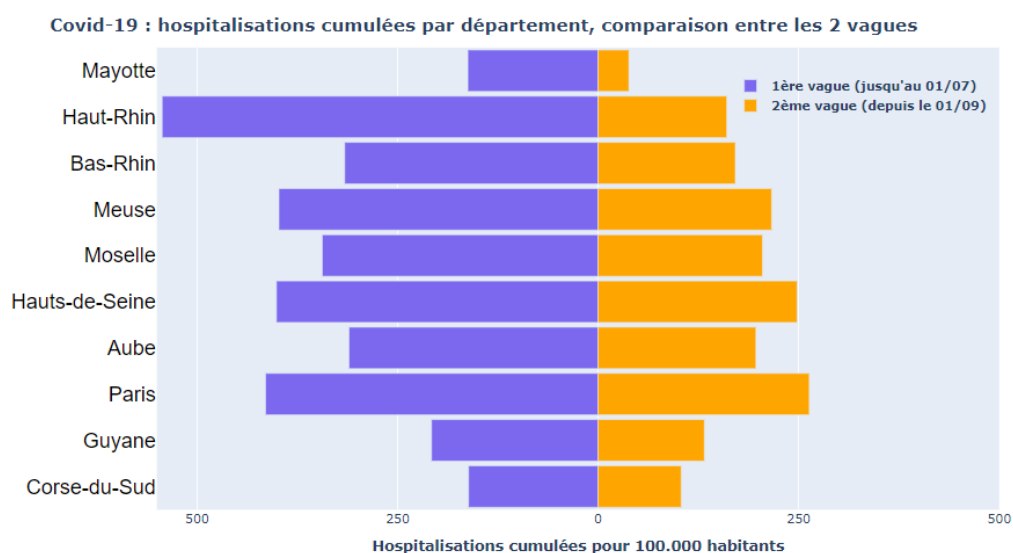


Départements ayant le taux d'incidence le plus faible - 14/12/2020 (7j glissants)

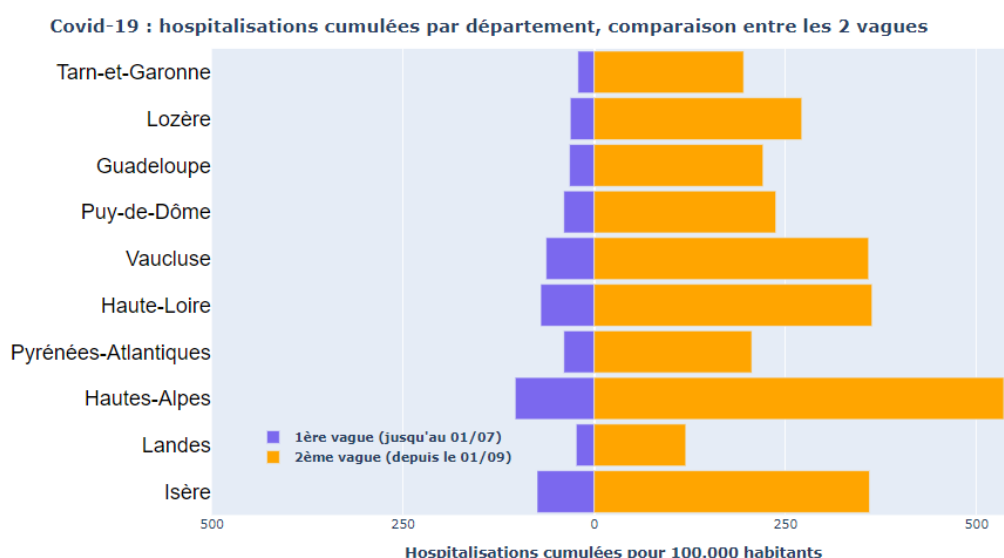


C – ANALYSE PAR DEPARTEMENT

2 – Disparités entre les deux vagues selon les départements

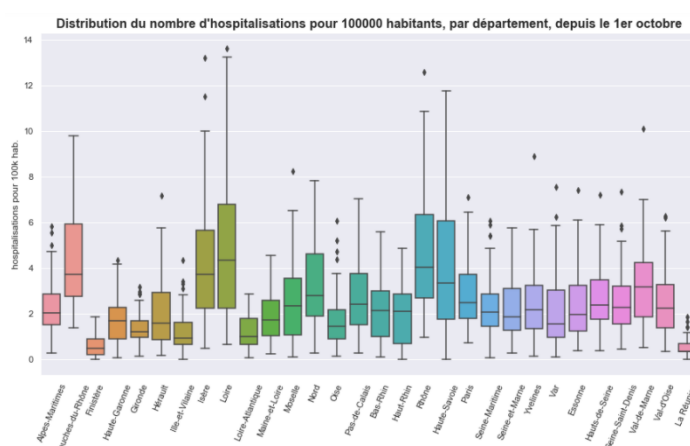
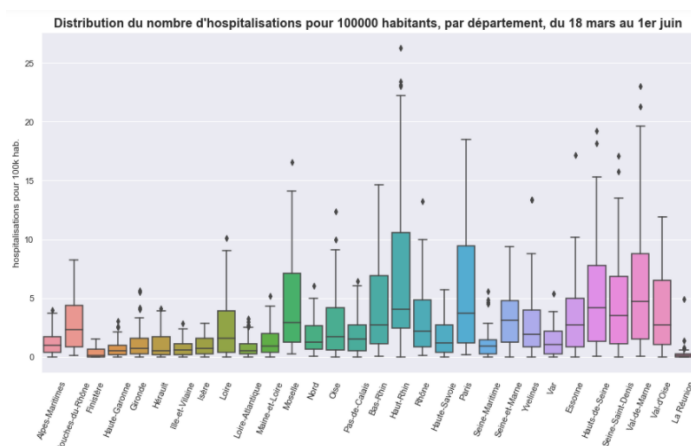


Certains départements ont été très fortement touchés par la première vague, et jusqu'ici relativement épargné par la seconde vague. Le Haut-Rhin, durement impacté durant le pic d'avril, compte près de cinq fois moins d'hospitalisations durant cette seconde vague. La ville de Paris est notamment moins impactée ces derniers mois, relativement au début de l'année.



A l'inverse, certains départements, n'ayant compté que très peu d'hospitalisations au printemps dernier, accusent un nombre d'hospitalisations bien plus élevé durant cette seconde vague. Le Tarn-Et-Garonne n'avait compté que 21 personnes hospitalisées pour 100,000 habitants au début de l'année. Depuis septembre, ce même département cumule plus de 360 patients pour 100,000 habitants admis à l'hôpital.

C – ANALYSE PAR DEPARTEMENT



La distribution des nouvelles hospitalisations quotidiennes, rapportées au nombre d'habitants, permet d'apprécier l'intensité de l'épidémie au sein d'un département. Le Haut-Rhin, durant la première vague, a dû faire face à un pic très violent avec quelques journées où les hospitalisations ont été extrêmement nombreuses. Les Hauts-De-Seine, malgré un nombre d'hospitalisations moyen similaire, n'ont pas eu à subir un tel afflux de patients au cours d'une seule journée.

Lors de la seconde vague, L'Isère, la Loire et le Rhône sont les départements qui ont subi ponctuellement les journées les plus difficiles, avec des pics supérieurs à 12 hospitalisations pour 100,000 habitants. Néanmoins ces valeurs restent bien inférieures aux plus dures journées de la première vague, où certains départements ont dépassé les 25 hospitalisations pour 100,000 habitants.

3 – Les hospitalisations hebdomadaires comme indicateur

Covid-19 : évolution par département du nombre de personnes hospitalisées sur 7 jours

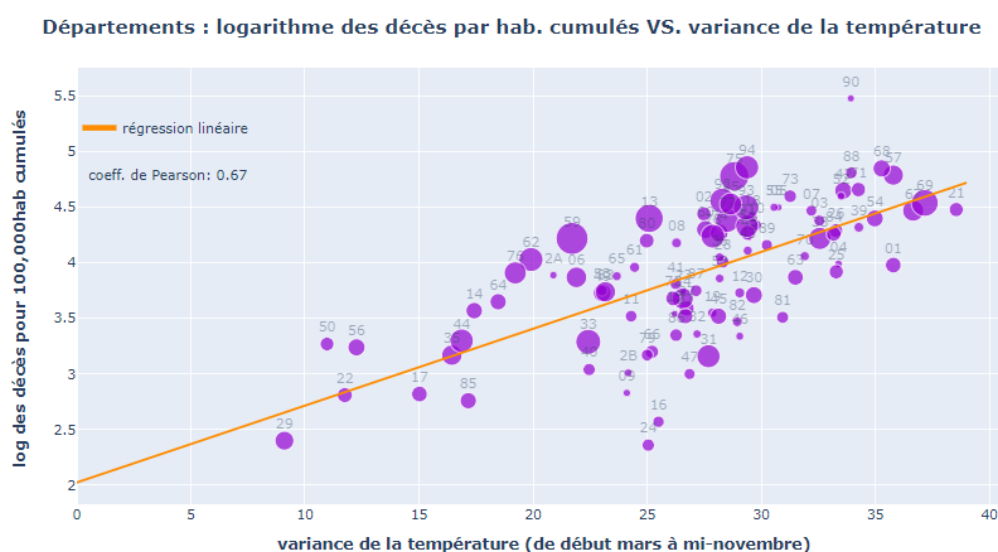


La différence entre les personnes hospitalisées sur sept jours et les personnes hospitalisées la semaine précédente permet de suivre l'évolution de l'épidémie au sein d'un département. Si cet indicateur est durablement positif, il est fort probable que le virus soit fortement présent au sein du département et que la situation se dégrade de plus en plus rapidement.

D – ETUDE DE CORRELATIONS

Nous cherchons ici à expliquer pourquoi certains départements, sur l'ensemble de l'année, ont été plus touchés que d'autres par l'épidémie de COVID-19. Nous étudions d'éventuelles corrélations entre le nombre de décès cumulés dans le département, rapporté à la population de ce dernier, et un certain nombre de variables explicatives.

1 – Variance de la température



Nous nous sommes également intéressés au facteur de la température pour expliquer la variance des décès. Nous étudions ici la variabilité de la température quotidienne au sein d'un département. La variable utilisée ici est la variance non biaisée de la température moyenne sur la période allant de début mars à la mi-novembre.

Il existe une corrélation significative entre cette variable et le logarithme du nombre de décès cumulés par habitant. Le coefficient de corrélation de Pearson est égal à 0.67. Cette valeur est relativement importante. Plus la variabilité de la température au sein d'un département est importante, plus il aura tendance à être durement touché par le virus.

On observe que les départements bretons, et plus généralement les départements bordant l'Atlantique subissent une très faible variance de température au fil de l'année. Ces départements sont également les plus épargnés par la pandémie, avec moins de 30 décès pour 100,000 habitants pour le Finistère ou encore les Côtes-d'Armor. Les départements les plus touchés, comme le Rhône ou le Haut-Rhin, sont parmi ceux marqués par une très forte variabilité de la température de jour en jour.

SOURCES

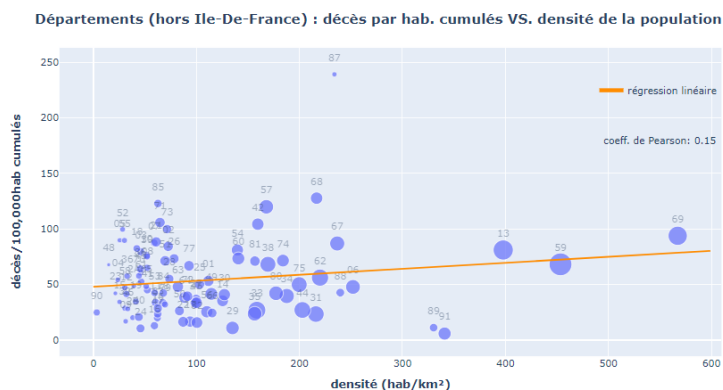
Les données démographiques et de température par département proviennent du site de l'INSEE

D – ETUDE DE CORRELATIONS

2 – Densité de la population

La concentration de la population au sein d'un département est un candidat potentiel pour expliquer la variance des décès. Un nombre d'habitants par kilomètre carré plus élevé pourrait intuitivement faciliter la transmission du virus, avec des contacts en moyenne plus nombreux entre les individus, notamment dans les commerces ou les transports.

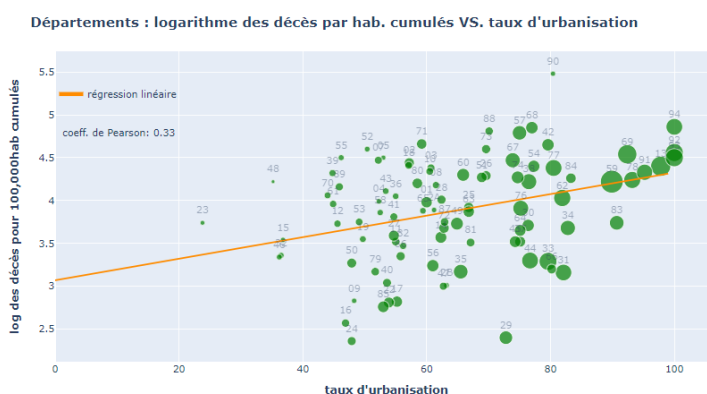
Néanmoins l'analyse bivariable de cette donnée avec le nombre de décès par habitants ne permet pas de détecter une corrélation significative. Sur l'ensemble de l'année, le coefficient de corrélation de Pearson est de 0,15.



3 – Taux d'urbanisation

Nous étudions enfin la relation entre le nombre de décès par habitant dans un département et le taux d'urbanisation dans ce dernier. Habiter dans une ville signifie bien souvent être en contact avec davantage d'individus qu'à la campagne. Les citadins se croisent dans les transports en commun, sur les lieux de travail, dans les restaurants et d'autres lieux publics, souvent en intérieur.

Nous mesurons une corrélation moyenne entre cette variable et le logarithme du nombre de décès cumulés, avec un coefficient de Pearson égal à 0,33. Le taux d'urbanisation n'a donc pas un impact inexistant sur la propagation de l'épidémie, mais on constate que les départements ayant le taux



d'urbanisation le plus élevé, comme les départements d'Ile-De-France, ne sont pas nécessairement les départements les plus touchés.

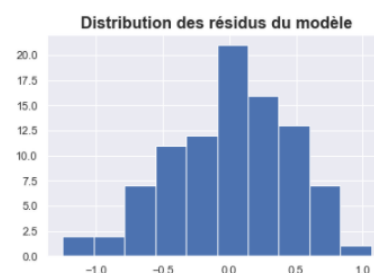
D – ETUDE DE CORRELATIONS

4 – Régression linéaire multiple

Nous tentons de modéliser le nombre de décès par habitant d'un département à l'aide d'une régression multiple, avec comme variables explicatives la variance de la température et le taux d'urbanisation.

OLS Regression Results						
=====						
Dep. Variable:	log_décès_p_hab	R-squared:	0.450			
Model:	OLS	Adj. R-squared:	0.438			
Method:	Least Squares	F-statistic:	36.48			
Date:	Fri, 18 Dec 2020	Prob (F-statistic):	2.68e-12			
Time:	11:38:55	Log-Likelihood:	-57.033			
No. Observations:	92	AIC:	120.1			
Df Residuals:	89	BIC:	127.6			
Df Model:	2					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

Intercept	1.7141	0.268	6.393	0.000	1.181	2.247
temp_var	0.0589	0.008	7.407	0.000	0.043	0.075
tx_urbanisation	0.0127	0.003	4.296	0.000	0.007	0.019
=====						
Omnibus:	2.584	Durbin-Watson:	1.844			
Prob(Omnibus):	0.275	Jarque-Bera (JB):	2.459			
Skew:	-0.395	Prob(JB):	0.292			
Kurtosis:	2.871	Cond. No.	394.			
=====						



Les coefficients du modèle obtenus sont tous significatifs, avec une p-value bien inférieure au seuil standard de 5%. Le test de Fisher est également satisfait, le modèle linéaire est donc ici pertinent. Le coefficient de la variable de température est le plus important (0,0589), et est positif, ce qui confirme l'analyse de corrélation réalisée plus tôt : la variance de la température est le meilleur facteur explicatif parmi les variables que nous avons testé.

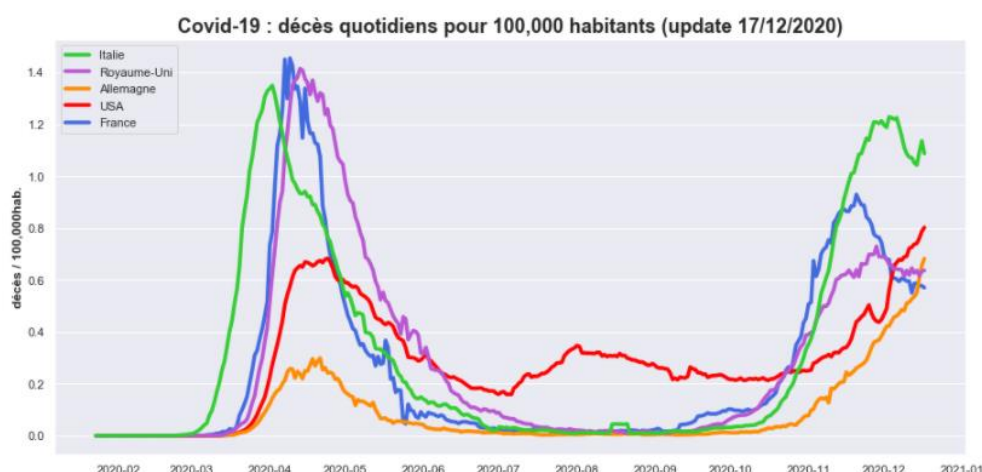
Le test de Jarque-Bera de normalité des résidu est lui aussi satisfait. Sa p-value est égale à 0,29, bien supérieur au seuil de 5%. On ne rejette donc pas l'hypothèse H_0 de normalité des résidus. Cette conclusion est confirmée par l'histogramme des résidus. La distribution de ces derniers est assez proche d'une distribution gaussienne pour valider les résultats du modèle.

Le R^2 de ce cette régression est de 0,45. On peut donc estimer que les variables choisies expliquent 45% de la variance des décès entre les différents départements français.

CONCLUSION

La deuxième vague est derrière nous, mais les chiffres ne diminuent plus. Les nouveaux cas sont de nouveaux en légère hausse et les hospitalisations stagnent. Nous projetons un maintien des chiffres sur ce plateau durant les deux dernières semaines de l'année.

Le solde sur sept jours des entrées en réanimations et celui des hospitalisations devront être scrutés après les fêtes, aussi bien au niveau national que départemental. Les taux d'incidences par département sont de nouveau sur une tendance haussière et le R effectif est repassé au-dessus de 1, signe d'un début de rebond.



L'évolution de la situation dans le reste du monde n'est pas bonne, puisqu'on observe un rebond des décès en Italie ou encore en Allemagne. Les États-Unis subissent également des nombres records de décès quotidiens.

Les prochaines semaines seront donc décisives, et dans l'attente de la vaccination de la population, la France reste sujette à une dégradation rapide de la situation et doit anticiper un possible rebond début 2021.