

Covid-19

L'odyssée d'une pandémie

Analyse préliminaire de la pandémie du coronavirus

Projet 8

Réaliser un projet libre dans le domaine professionnel qui vous intéresse

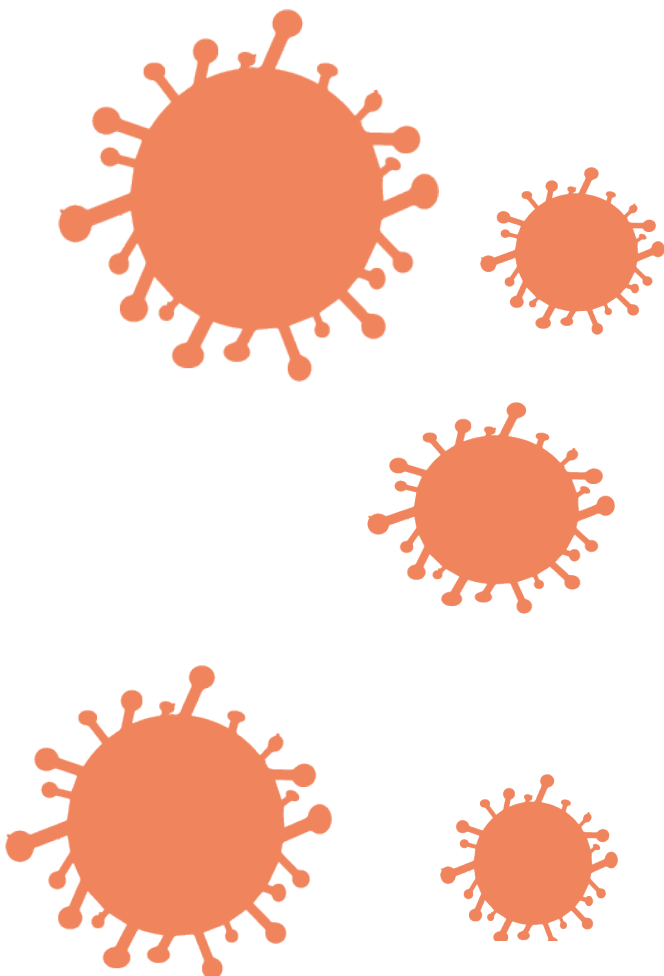


Table des matières

1. Qu'est-ce que le coronavirus?	2
1.2 Les symptômes	3
1.3. Contexte mondial	4
2. Les données	5
3. Chronologie de la pandémie	7
3.1 Le monde	7
3.2 Les pays les plus touchés	8
4. A la recherche d'un modèle de prediction	10
4.1 Séries Temporelles	10
4.2 Machine learning	12
5. La France	13
6. Conclusions	14
7. Références	15
8. Annexe	16

Avril 2020

A l'évidence des changements soudain dans les paradigmes mondiaux à cause de la pandémie du nouveau coronavirus, Covid-19 je trouve plus que pertinent de mettre mon savoir-faire comme Data Analyst pour enrichir la recherche à fin de comprendre le fonctionnement de transmission de la maladie.

Les objectif de ce travail seront :

1. Évolution dans le nombre de cas touchés au fil du temps
2. Évolution des cas au fil du temps au niveau des pays
3. Créer un modèle prédictif performant

1 Qu'est-ce que le coronavirus?

Les coronavirus forment une vaste famille de virus qui peuvent être pathogènes chez l'homme et chez l'animal. On sait que, chez l'être humain, plusieurs coronavirus peuvent entraîner des infections respiratoires dont les manifestations vont du simple rhume à des maladies plus graves comme le syndrome respiratoire du Moyen-Orient (MERS) et le syndrome respiratoire aigu sévère (SRAS). (figure 1) Le dernier coronavirus qui a été découvert est responsable de la maladie à coronavirus 2019 (COVID-19). Ce nouveau virus



Figure 1-Covid-19

et cette maladie étaient inconnus avant l'apparition de la flambée à Wuhan (Chine) en décembre 2019. Le 20 janvier 2020, le point focal national RSI de la République de Corée a signalé le premier cas de nouveau coronavirus en République de Corée. Il s'agit d'une femme de 35 ans, de nationalité chinoise, résidant à Wuhan (province du Hubei, Chine).

La patiente a présenté les symptômes suivants : fièvre, frissons et douleurs musculaires, le 18 janvier, alors qu'elle se trouvait à Wuhan. Elle s'est rendue dans un hôpital local à Wuhan et le diagnostic posé initialement a été celui d'un rhume. Le 19 janvier, à son arrivée à l'aéroport international d'Incheon (Corée du Sud), la patiente avait de la fièvre (38,3°C). Elle a été transférée dans un hôpital d'isolement désigné à cet effet pour être soumise à des tests de dépistage et bénéficier d'un traitement. Les tests ont donné

un résultat positif à la RT-PCR (amplification génique après transcription inverse) pour les coronavirus; le séquençage effectué ensuite, le 20 janvier, par les *Centers for Disease Control and Prevention* de Corée (KCDC) ont permis de confirmer que les résultats étaient positifs pour le nouveau coronavirus (2019-nCoV). Au moment du dépistage, la patiente présentait les symptômes suivants : frissons, écoulement nasal et douleurs musculaires.

La patiente a indiqué ne s'être rendue sur aucun marché, y compris le marché aux poissons de Huanan; elle n'a pas non plus eu de contact connu avec des cas confirmés de 2019-nCoV ou des animaux sauvages dans la ville de Wuhan.

La patiente est actuellement en isolement et reçoit un traitement; son état est stable.

A partir du 20 janvier la propagation du virus au niveau mondiale se fait très rapidement.

Le 11 mars 2020 l'Organisation mondiale de la santé (OMS) qualifie l'épidémie du Covid-19 de « pandémie »

1.2 Les symptômes

Les symptômes les plus courants de la COVID-19 sont la fièvre, la fatigue et une toux sèche. Certains patients présentent des douleurs, une congestion nasale, un écoulement nasal, des maux de gorge ou une diarrhée (figure 2). Ces symptômes sont généralement bénins et apparaissent de manière progressive. Cer-

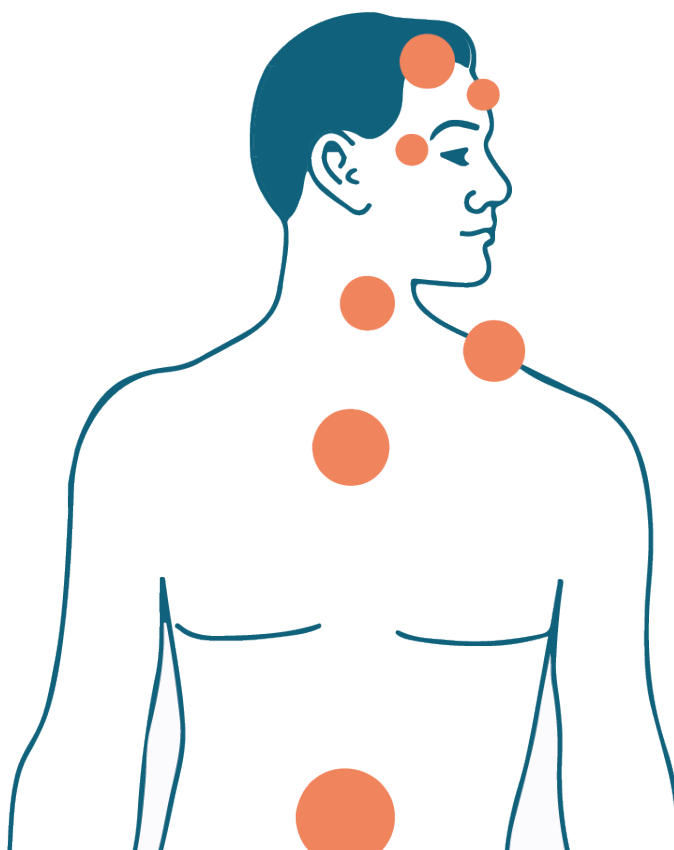


Figure 2. Covid-19 symptômes

Les symptômes en chiffres :

Fièvre :	87,9 %
Toux sèche :	67,7 %
Fatigue :	38,1 %
Expectorations :	33,4 %
Dyspnée :	18,6 %
Douleurs musculaires ou articulaires :	14,8 %
Maux de gorge :	13,9 %
Céphalées :	13,6 %
Frissons :	11,4 %
Nausée et vomissements :	5,0 %
Congestions nasales :	4,8 %
Diarrhées :	3,7 %
Hémoptysie :	0,9 %
Congestion conjonctivale :	0,8 %

taines personnes, bien qu'infectées, ne présentent aucun symptôme et se sentent bien. La plupart (environ 80 %) des personnes guérissent sans avoir besoin de traitement particulier. Environ une personne sur six contractant la maladie présente des symptômes plus graves, notamment une dyspnée. Les personnes âgées et celles qui ont d'autres problèmes de santé (hypertension artérielle, problèmes cardiaques ou diabète) ont plus de risques de présenter des symptômes graves. Toute personne qui a de la fièvre, qui tousse et qui a des difficultés à respirer doit consulter un médecin.

La fièvre peut apparaître plusieurs jours après la toux.

De l'anosmie et des pertes du goût peuvent également survenir, notamment chez les plus jeunes.

De nombreuses personnes porteuses du virus ne présentent pas de symptômes, ou présentent des symptômes légers (petite toux, fièvre) sans détresse respiratoire d'où aussi sa propagation si rapide.

La maladie du COVID-19 est transmise par des personnes porteuses du virus. La maladie peut se transmettre d'une personne à l'autre par le biais de gouttelettes respiratoires expulsées par le nez ou par la bouche lorsqu'une personne tousse ou éternue. Ces gouttelettes peuvent se retrouver sur des objets ou des surfaces autour de la personne en question. On peut alors contracter la COVID-19 si on touche ces objets ou ces surfaces et si on se touche ensuite les yeux, le nez ou la bouche. Il est également possible de contracter la COVID-19 en inhalant des gouttelettes d'une personne malade qui vient de tousser ou d'éternuer. C'est pourquoi il est important de se tenir à plus d'un mètre d'une personne malade.

L'OMS examine les travaux de recherche en cours sur la manière dont la COVID-19 se propage et elle continuera à communiquer les résultats actualisés.

1.3 Contexte mondial

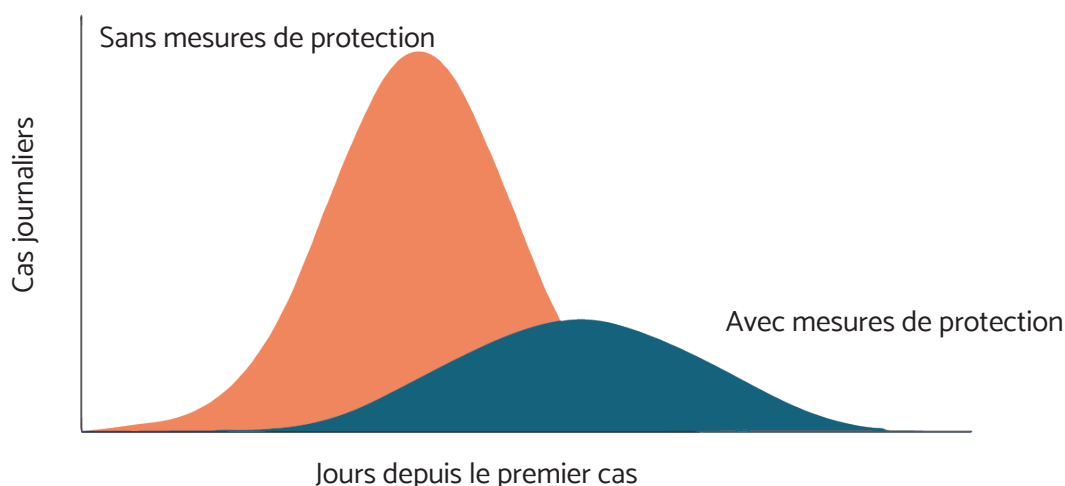


Figure 3 -Représentation de la courbe de personnes contaminées

A l'heure actuelle des mesures strictes pour ralentir la propagation du virus sont mises en place par les différents pays au niveau mondial. Ainsi, l'objectif des mesures politiques prises par les États (interdiction des rassemblements, mise en quarantaine, fermeture des écoles, transports, etc.) est "d'aplatir la courbe" de l'épidémie. Autrement dit, ralentir le rythme des nouvelles contaminations de façon à étaler dans le temps le nombre de cas nécessitant une hospitalisation.

Le graphique ci-dessus (figure 3) est réalisé à partir d'un schéma des Centres américains de contrôle des maladies à propos des épidémies de grippe. Il montre qu'en l'absence de mesures de protection ou de confinement, le nombre de cas augmente rapidement et dépasse les capacités de prise en charge des patients sévèrement atteints à l'hôpital (ligne en pointillés). Les mesures individuelles ou collectives permettent de retarder le pic épidémique pour étaler dans le temps le nombre de cas d'infection. L'épidémie n'est pas empêchée, mais suffisamment ralentie pour permettre aux hôpitaux de s'occuper correctement des malades. Plus le nombre de cas augmente rapidement, plus les mesures pour limiter les nouvelles contaminations deviennent nécessaires.

2 Les données

Source kaggle :

<https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset#novel-corona-virus-2019-dataset.csv>

Les données sont mises à jour, toutes les 24h. Nous sommes encore au milieu de la pandémie donc les analyses et la récolte des données n'ont pas assez de recul. Mais néanmoins, nous pouvons faire une analyse exploratoire des données et essayer de sortir quelques prédictions. Bien sûr, ces prédictions resteront à vérifier a posteriori. Nous avons une excellente opportunité pour tester nos algorithmes de prédiction et nos façons de récolter les données.

Tout au long de ce rapport nous allons travailler avec un dataframe avec cinq variables. Trois d'elles sont des variables quantitatives, les autres deux sont une variable temporelle et une variable catégorielle.

Voici le data frame :

DATE	COUNTRY	CONFIRMED	DEATHS	RECOVERED
du 22 janvier au 04 avril 2020	218 pays	nombre de cas confirmés par pays	nombre de décès à cause du coronavirus	nombre de cas de guérison

La variable «confirmed» réunit les chiffres journaliers des personnes testées positives au coronavirus par pays depuis que le premier cas a été enregistré . On doute que ses chiffres représentent le vrai nombre de personnes infectées car les tests ne se réalisent qu'une fois que la personne a des symptômes.

La variable «deaths» contient les chiffres journaliers des personnes décédées par complications respiratoires à cause du Covid-19. Et la variable «recovered» réunit les chiffres journaliers des personnes guéries. On trouve ces chiffres par pays, par région et par date.

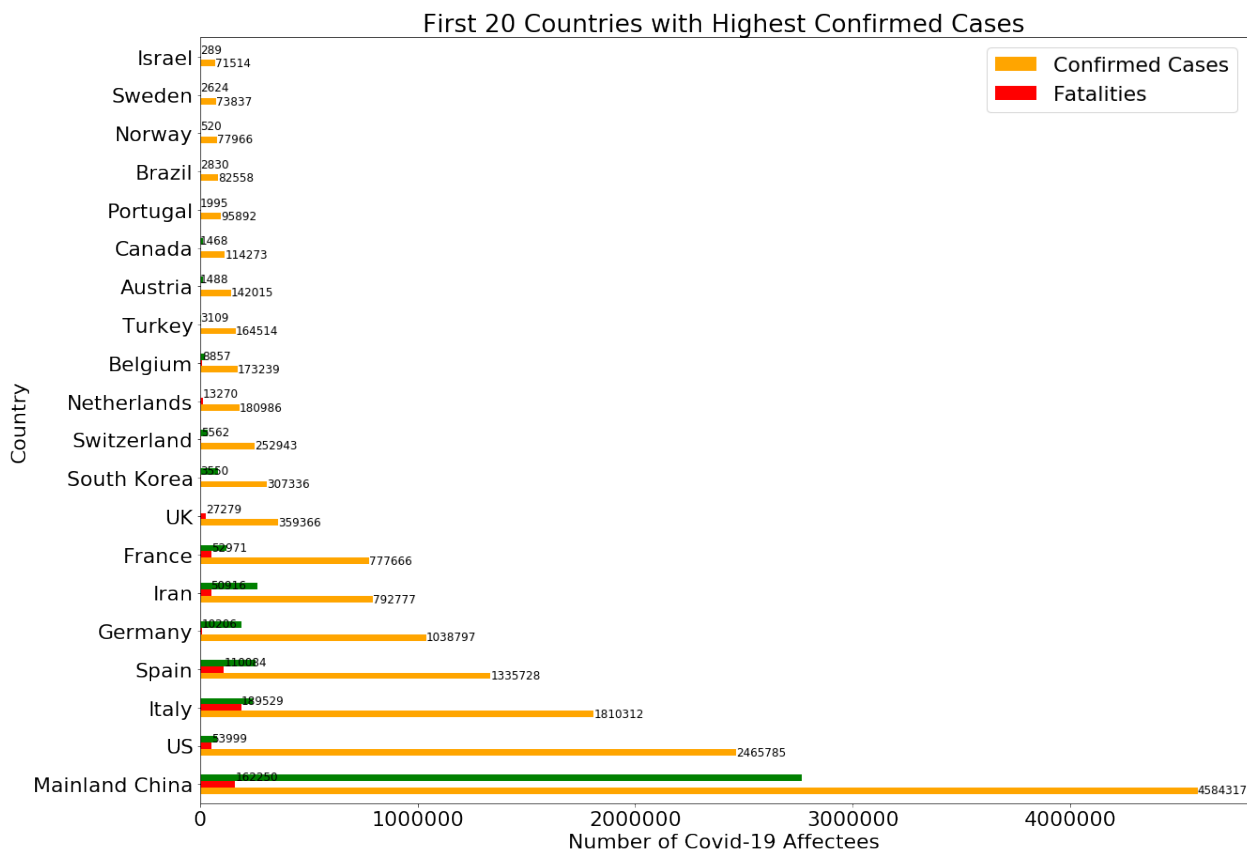


Figure 4- Les 20 pays les plus touchés au 04-04-2020

A l'œil nu nous pouvons voir que les trois courbes suivent la même progression au fil du temps (figure 5). Elles ont un départ lent jusqu'à la deuxième semaine du mois de mars où les chiffres montent très rapidement. Par contre, comme nous voyons sur le graphique ci-dessus le pourcentage de mortalité change par rapport aux pays. Par exemple, la Chine est le pays avec les plus de cas confirmés mais n'a pas un taux de mortalité élevé. (figure 4).

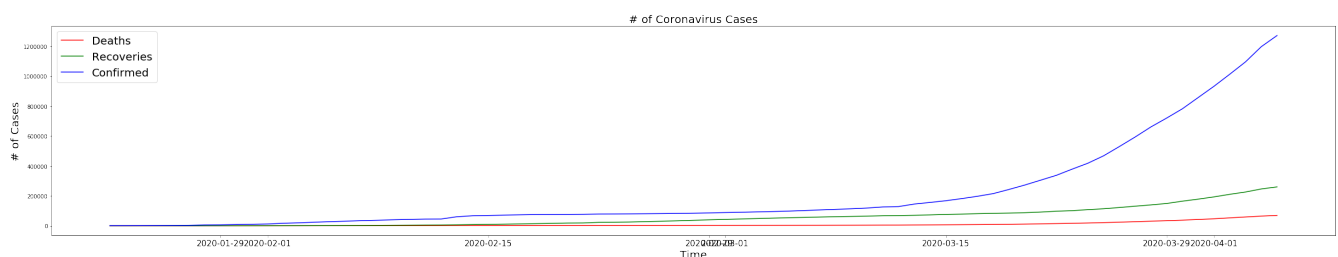


Figure 5 -Representation de l'évolution des trois variables au niveau mondiale

3 Chronologie de la pandémie

Évolution dans le nombre de cas touchés au fil du temps

Nous allons voir à partir des graphiques que la maladie a, clairement, commencé à se propager en Chine, et nous allons, également, observer que l'Italie et les États-Unis suivent la tendance de la Chine et ont des chiffres élevés.

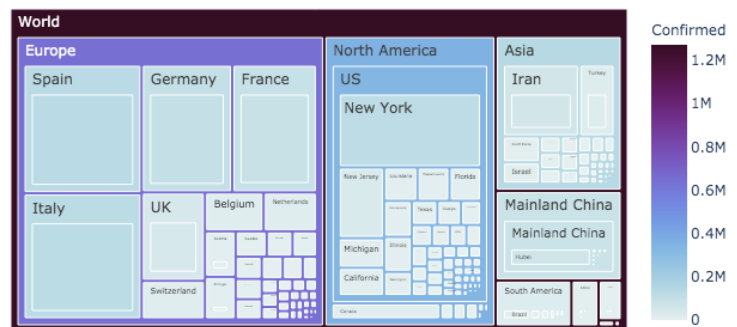
3.1 Le monde

Des médecins chinois donnent l'alerte sur un nouveau virus inconnu qu'ils nomment 2019n-CoV et qui serait apparu chez des personnes travaillant au marché de gros de fruits de mer de Huanan à Wuhan dans la province de Hubei en Chine centrale, dans lequel des animaux sauvages vivants sont entreposés et vendus. Ils sont très vite mis sous silence par le gouvernement chinois qui espère régler cette affaire en interne, mais il s'avère que ce virus ressemble à l'épidémie du SRAS ayant surgi en 2002-2003 en Chine également et qui avait été assez meurtrière.

Le virus se propage rapidement sur tout le territoire chinois durant les deux mois suivants : le 30 janvier, la région autonome du Tibet est la dernière région infectée.

.Dans les graphiques à gauche on peut voir les cas par région et par pays. C'est très significatif de voir qu'il y a des villes très affectées comme par exemple New York..

Current share of Worldwide COVID19 Cases



Current share of Worldwide COVID19 Deaths

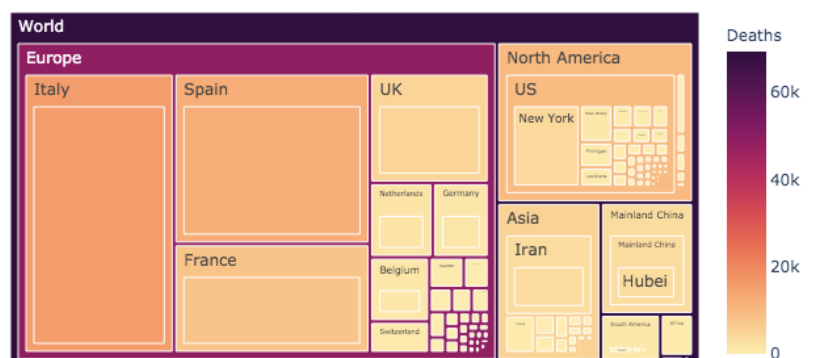


Figure 6 et 7 - Pourcentage actuel de personnes contaminées par le Covid.19 dans le monde

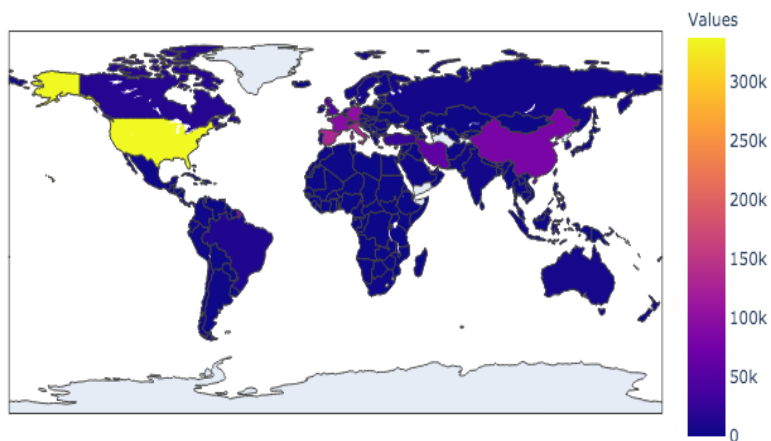


Figure 8 -Situation mondiale au 06-04-2020

30-12-19 27 personnes testées positive au Covid-19 à Wuhan
 11-01-20 Première personne décédée
 13-01-20 Premier cas hors Chine
 31-01-20 10.000 cas et 170 décédés
 07-02-20 500 personnes décédées
 10-02-20 1.000 personnes décédées
 23-02-20 Première personne décédée en Italie
 29-02-20 87.000 cas de coronavirus au monde et le premier cas de décès en Etats-Unis
 01-03-20 90.000 cas confirmés au monde et 3.000 personnes décédées
 03-03-20 Premier cas de décès en Espagne
 09-03-20 9.000 cas et 465 décès

11-03-20 L'OMS déclare le Covid_19 comme une pandémie
 13-03-20 4.700 personnes décédées
 16-03-20 6.500 personnes décédées
 19-03-20 Record de 475 décès en 24h en Italie
 01-04-20 1.000.000 de cas confirmés de Covid_19 au monde
 14-04-20. 2.000.000 de cas confirmés de Covid_19 au monde. Les Etats Unis deviennent le pays le plus touché avec 26.100 personnes décédées et un record de 2200 décès en 24h (voir dans l'annexe figure 3)

Dans les premiers jours de janvier l'OMS finit par être alertée sur le virus. Les gouvernements, dont celui de Chine, n'y font pas trop attention. Plusieurs aéroports mettent en place un dépistage à l'arrivée des avions provenant de Chine. Celui-ci s'effectue par contrôle de la température frontale des voyageurs avec un thermomètre électronique. Cependant, ces mesures ne peuvent enrayer la propagation de l'épidémie en dehors du territoire chinois.

Le 13 mars l'OMS déclare que désormais « l'Europe est à l'épicentre de la pandémie ».

Fin mars, le Royaume-Uni avait déjà dépassé la barre des 1 000 morts et l'Italie le nombre de 10 000 morts. Il y a deux points de fluctuation sur la période de Coronavirus, un entre le 10 février et le 15 février et l'autre entre le 10 mars et le 15 mars, ce qui est exactement une période d'un mois. Donc, il pourrait suivre un modèle cyclique, nous allons essayer d'explorer à partir de cette perspective.

3.2 Les pays les plus touchés

Nous allons démontrer que l'évolution de la pandémie est directement liée aux mesures de contention de la propagation du virus prise par chaque pays.

Aussi, notamment, au moment où ces mesures sont mis en place. Avant toute analyse, il faut préciser que la façon de compter les cas de personnes confirmées porteuses de coronavirus et les cas de personnes décédés à cause du virus n'est pas le même pour tous les pays.

Prenons les cas de quatre pays emblématiques :

l'Allemagne, la Chine, l'Italie et la France.

L'Allemagne avec un taux de mortalité de 1% est le pays avec le plus faible taux .(figure 12)

Dans l'autre extrême nous allons trouver l'Italie avec un taux de mortalité de 9,5%. (figure 9)

Entre ces deux extrêmes, nous trouverons la France avec 6,4% et la Chine avec le 3,4%. (figure 10 et 11)

Prenons d'abord la Chine, ils se sont placés en quarantaine le 22 janvier avec un total de 547 cas des personnes confirmées contaminées par le covid-19. Avec ces mesures prématurées ils ont réussi à maintenir les nombres de cas de personnes contaminées presque à plat depuis le 20 février 2020.

En comparaison, l'Italie a commencé la quarantaine le 9 mars 2020 avec 9172 cas de personnes confirmées malades de coronavirus. A l'heure actuelle, le nombre de cas de personnes malades augmente de façon exponentielle et les cas de décès aussi. C'est ce qu'explique le pourcentage très élevé de mortalité dans ce pays.

Au 9 mars la Chine avait déjà réussi à contenir la propagation du virus, contrairement au reste du monde

Mortalité dans quatre pays extrêmes

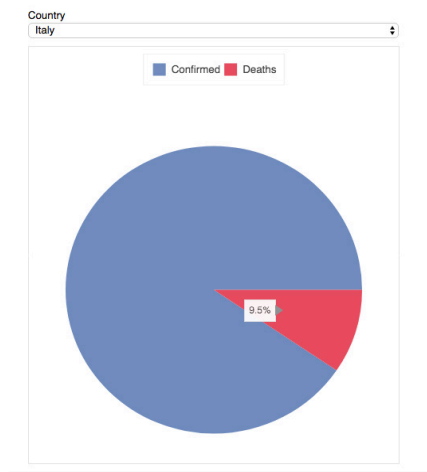


Figure 9-La mortalité en Italie

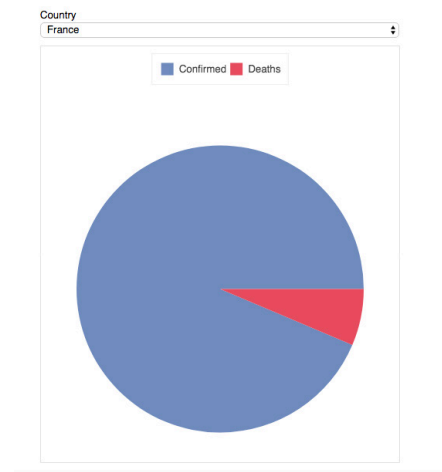


Figure 10-La mortalité en France

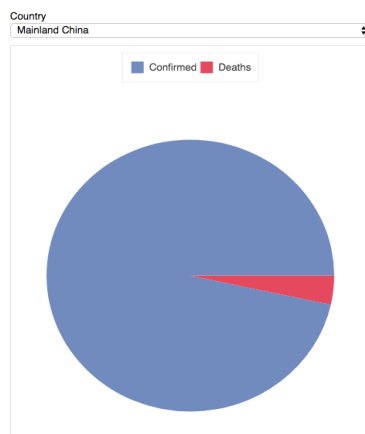


Figure 11 -La mortalité en Chine

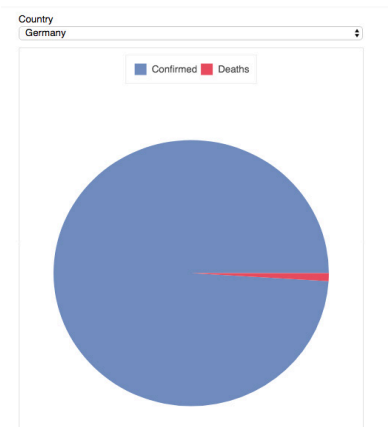


Figure 12 -La mortalité en Allemagne

qui voit ses cas monter très rapidement.

A partir de mi mars nous allons voir comme (très rapidement) l'Italie, la France et notamment les Etats-Unis arrivent à dépasser la Chine en nombre de cas confirmés de personnes contaminées. Et le taux de mortalité s'élève à tout vitesse.

En observant les cas confirmés et les cas de décès au fil du temps, l'Italie est sur le point de répéter le chemin de la Chine, tandis que le taux de mortalité en Italie semble être le pire de tous et il suit une tendance croissante.

On a vu l'impact des mesures de contention sur la mortalité du virus. On constate donc une forte corrélation entre ces deux variables. (voir dans l'annexe figure 1 et 2)

4 A la recherche d'un modèle de prédiction

Créer un modèle prédictif performant

4.1 Séries Temporelles

Une série temporelle, ou série chronologique, est une suite de valeurs numériques représentant l'évolution d'une quantité spécifique au cours du temps. De telles suites de variables aléatoires peuvent être exprimées mathématiquement afin d'en analyser le comportement, généralement pour comprendre son évolution passée et pour en prévoir le comportement futur.

Holt Winters

Ce modèle est une extension du lissage exponentiel de Holt qui capture la saisonnalité. Cette méthode génère des valeurs lissées de façon exponentielle pour le niveau, la tendance et l'ajustement saisonnier de la prévision. Cette méthode additive saisonnière ajoute le facteur de saisonnalité à la prévision de tendances, ce qui aboutit à la prévision additive de Holt-Winters.

Cette méthode convient le mieux aux données avec une tendance et une saisonnalité qui n'augmente pas au fil du temps. Elle aboutit à une prévision en courbe qui présente les modifications saisonnières des données.

Pourtant on a vu que notre série temporelle est multiplicative. Pour cela avant de mettre en place notre modèle de Holt Winters on a transformé la série en additive avec la fonction log.

La prediction de cas confirmés

Nous allons diviser nos données pour pouvoir entrainer notre modèle en deux parties. 80% pour l'entraînement et 20 % pour le test.

Nous pouvons voir sur le graphique les résultats de la prediction du modèle et son acuité.(figure 13)

Il est clair sur le graphique que la prediction suit parfaitement la même ligne du set de validation.

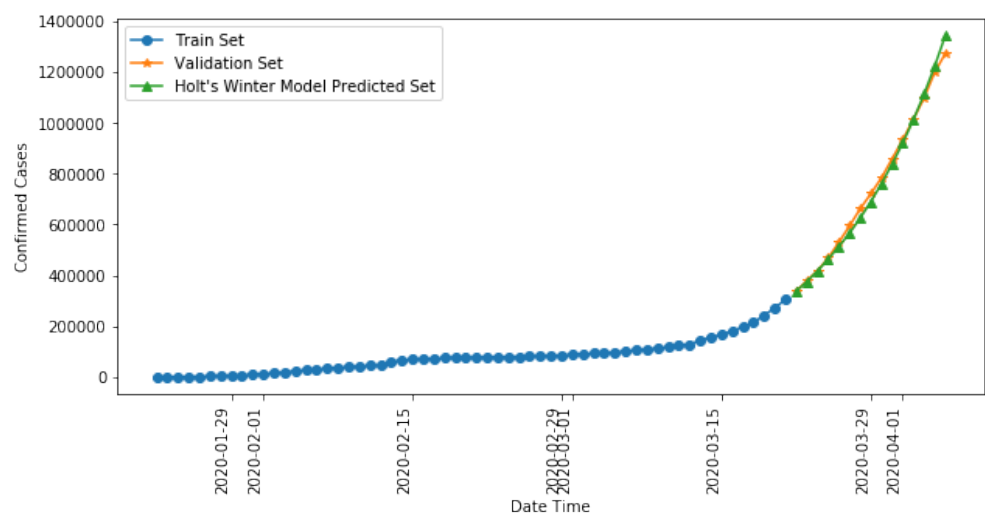


Figure 13- Resultat de la prediction du modèle Holt Winters

ARIMA

Nous allons utiliser le modèle ARIMA, pour essayer de prédire l'évolution des cas confirmés de coronavirus. Les modèles de la famille ARIMA permettent de représenter sous une forme succincte certains phénomènes variant avec le temps, et de faire des prévisions pour les valeurs futures du phénomène, avec un intervalle de confiance autour des prévisions.

Les processus autorégressifs supposent que chaque point peut être prédit par la somme pondérée d'un ensemble de points précédents, plus un terme aléatoire d'erreur.

Le processus d'intégration suppose que chaque point présente une différence constante avec le point précédent.

Les processus de moyenne mobile supposent que chaque point est fonction des erreurs entachant les points précédant, plus sa propre erreur.

L'écriture mathématique des modèles ARIMA suppose soit X_t une série chronologique de moyenne μ . Le modèle ARIMA(p,d,q), a les paramètres suivants :

p est l'ordre de la partie autorégressive du modèle.

q est l'ordre de la partie moyenne mobile du modèle.

d est l'ordre de différentiation du modèle.

Modèle pour le cas des décès

$$p = 1 \quad q = 1 \quad d = 1$$

Pour évaluer le résultat du modèle il faut déterminer :

- si chaque terme dans le modèle est significatif
- l'ajustement du modèle aux données
- si notre modèle respecte l'hypothèse de l'analyse

Notre modèle réunit ces trois conditions et nous pouvons le voir dans le graphique avec le résultat de la prédiction (figure 14)

Ici on voit comment au départ le modèle arrive à prédire très précisément. Par contre, à la fin la prédiction s'écarte du set de validation.

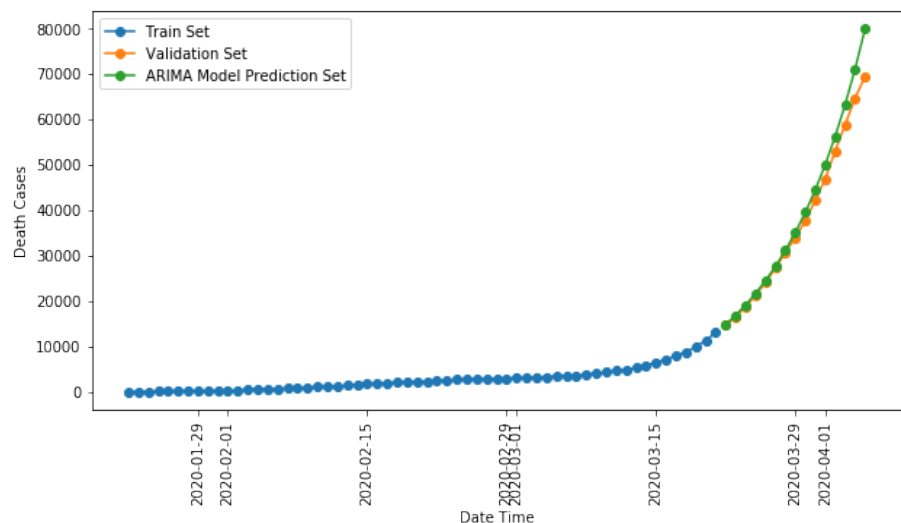
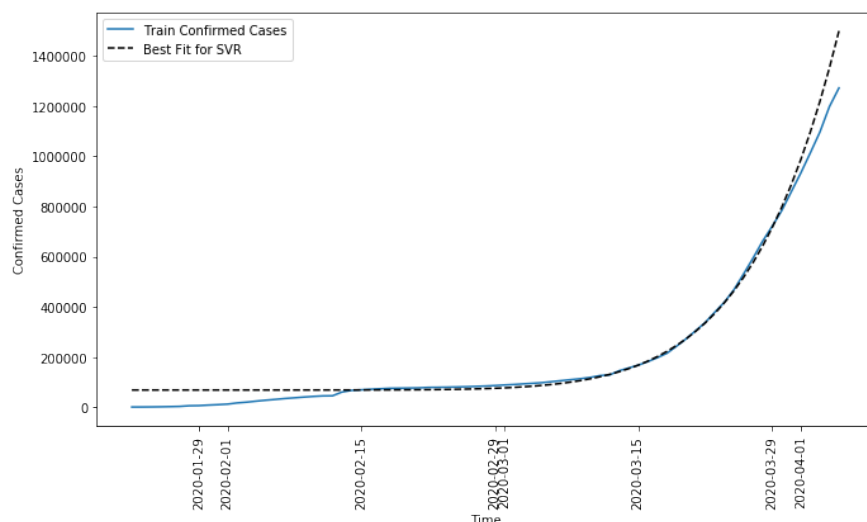


Figure 14 - Résultat de la prédiction du modèle ARIMA

4.2 Machine learnig

A la lumière des résultats des prédictions des séries temporelles, nous allons essayer des méthodes de machine Learning pour arriver à avoir un modèle prédictif plus performant et voir s'il tient mieux que les modèles de serie temporelles.

L'apprentissage automatique est un champ d'étude de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'« apprendre » à partir de



données, c'est-à-dire d'améliorer leurs performances à résoudre des tâches sans être explicitement programmés pour chacune. Nous, on va se servir d'une des méthodes de Machine Learning, appelée Support Vecteur Machine pour construire un modèle afin de prédire la progression de la pandémie.

Les SVMs sont une famille d'algorithmes d'apprentissage automa-

Figure 15 -Résultat de la prediction du modèle SVM pour les cas confirmés

tique qui permettent de résoudre des problèmes tant de classification que de régression ou de détection d'anomalie. Ils sont connus pour leurs solides garanties théoriques, leur grande flexibilité ainsi que leur simplicité d'utilisation même sans grande connaissance de data mining.

La prédiction de cas confirmés

Nous allons entrainer notre algorithme de Svm de la même façon qu'avec les modèles de série temporelle, c'est-à-dire, avec le 80% des données.

Le résultat est très intéressant mais moins performant que notre modèle d'Holt Winters pour le cas confirmé. Nous sont qu'à la fin la prédiction s'éloigne de la courbe du set de validation. (figure 15)

La prédiction des cas de décès

Le modèle SVM marche beaucoup mieux avec les cas de décès qu'avec les cas confirmés de covid-19. Notre modèle de prévision possède déjà de bonnes propriétés. Il faudrait voir s'il continue à l'être à fur et mesure que le temps passe et dès que des nouvelles données sont générées.

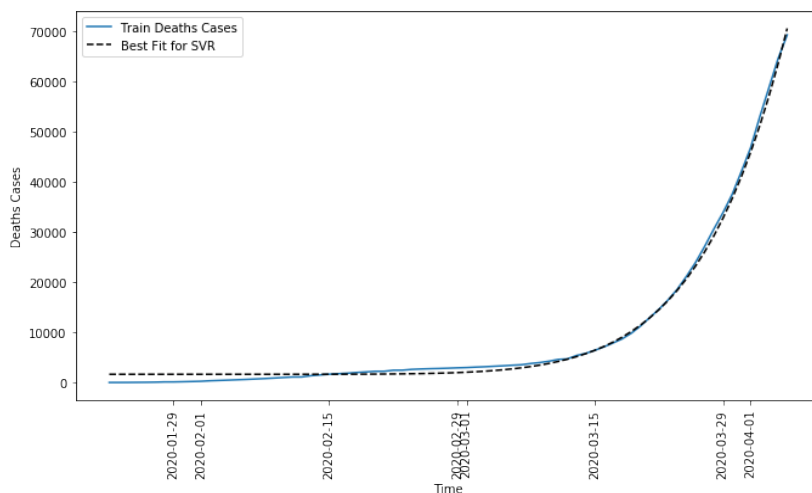


Figure 16 -Résultat de la prediction du modèle SVM pour les décès

5 La France

L'état des lieux

Au moment d'écrire ce rapport la situation sanitaire de la France est inquiétant. Le nombre de cas de décès monte trop rapidement. (figure 17)

Les journaux décrivent des services de réanimation au bord de la rupture. Pour atteindre l'objectif fixé par l'Exécutif de 14 500 lits de réanimation, les interlocuteurs du Figaro se montrent cependant confiants. Le président de la Société de réanimation de langue française (SRLF) lance pour sa part un appel aux industriels : ils doivent aider à produire en masse des respirateurs. La presse se fait l'écho de ces initiatives. Les journaux détaillent également la « course aux masques », pointant une concurrence entre les pays pour les commandes. Le transfert de malades est également relaté. Dans les Ehpad, Le Figaro note que les annonces du gouvernement reçoivent plutôt un bon accueil. Enfin, le traitement à la chloroquine et la figure de D. RAOULT sont toujours présents, comme la prudence du gouvernement à ce sujet. En Une, La Croix a recueilli le témoignage de patients guéris.

En addition à la situation sanitaire il y a aussi l'impact économique que cette pandémie est en train de produire. Les Echos et L'Opinion s'inquiètent de l'ampleur réelle de l'effort budgétaire, encore difficile à chiffrer. L'Opinion se demande qui paiera et craint une « explosion » des impôts. « Une fois la crise sanitaire endiguée, l'objectif sera de faire repartir l'économie avec un plan de relance. On va plutôt chercher à donner de l'air aux ménages et aux entreprises », assure l'entourage de G. DARMANIN, après les propos rassurants du ministre au Parisien. Même son de cloche du côté de la majorité. Malgré les mesures de soutien de l'Etat, nombre de dirigeants s'inquiètent pour le paiement de leurs charges courantes, souligne par ailleurs Les Echos, décrivant des entreprises en manque de trésorerie. Le président de la CCI Nouvelle-Aquitaine en témoigne dans un entretien. Il juge que « le plan de soutien à l'économie mis en place par le gouvernement est déjà très important », mais il appelle à des messages « plus clairs, concordants et précis afin de rassurer les entrepreneurs ».

France through Corona over a period of Time

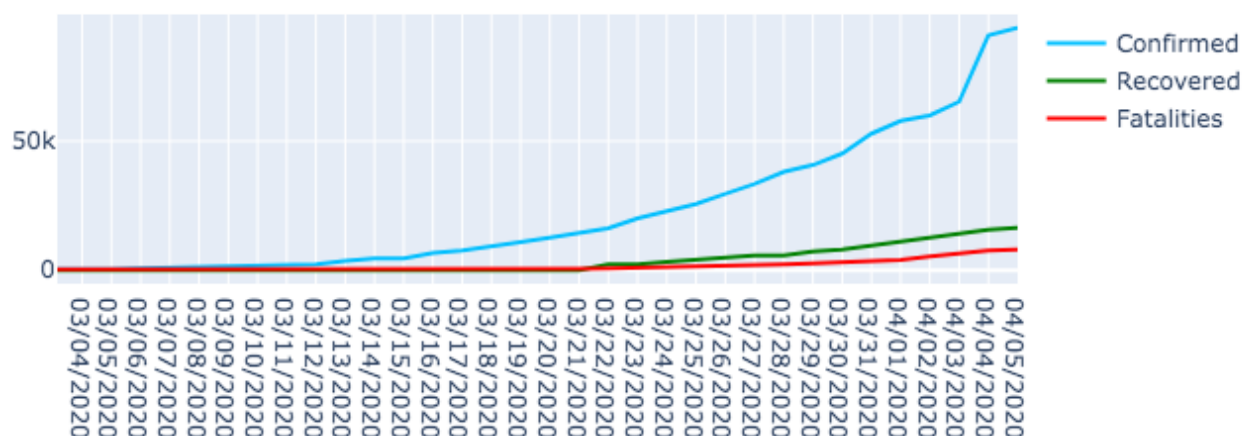
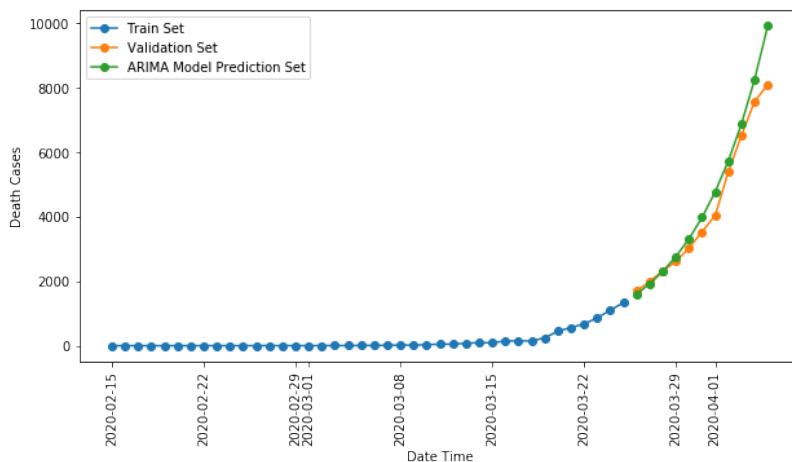


Figure 17 - La France : les chiffres des cas confirmés, décédés et guéris dans le temps

Appliquer les modèles de prédiction de cas de décès à la France

Nous allons voir que la série temporelle de la France suit la même progression de la courbe que la série temporelle avec les chiffres mondiaux. Le modèle Arima donne des résultats assez pertinents avec les données françaises. (figure 18)

Les modèles en plus de faire des prédictions aident aussi à mieux comprendre le virus. En comparant les prédictions aux statistiques du terrain, nous pouvons repérer les paramètres qui expliquent les éventuelles différences entre l'évolution des chiffres au niveau mondial et l'évolution au niveau de la France. Ainsi, nous pouvons évaluer si les mesures mises en places pour aplatis la propagation du virus sont réellement efficaces.



Sur le graphique à gauche nous pouvons voir comme le modèle Arima arrive à donner des résultats de prédiction très intéressants. (figure 18)

Le graphique ci-dessous montre les résultats du modèle de machine learning SVM (figure 3)

Au 4 avril 2020 nous sommes encore dans un rythme de propagation exponentielle et pour le moment nous n'avons pas assez de recul pour arriver à avoir un modèle

Figure 18 - Résultat de la prédiction du modèle Arima pour les décès en France

avec une acuité plus élevée. Néanmoins, nous arrivons à avoir des prédictions qui ne sont pas très éloignées de la réalité. Il reste à voir comment nous allons pouvoir entraîner nos modèles au fur à mesure que le temps passe et que les données seront plus nombreuses.

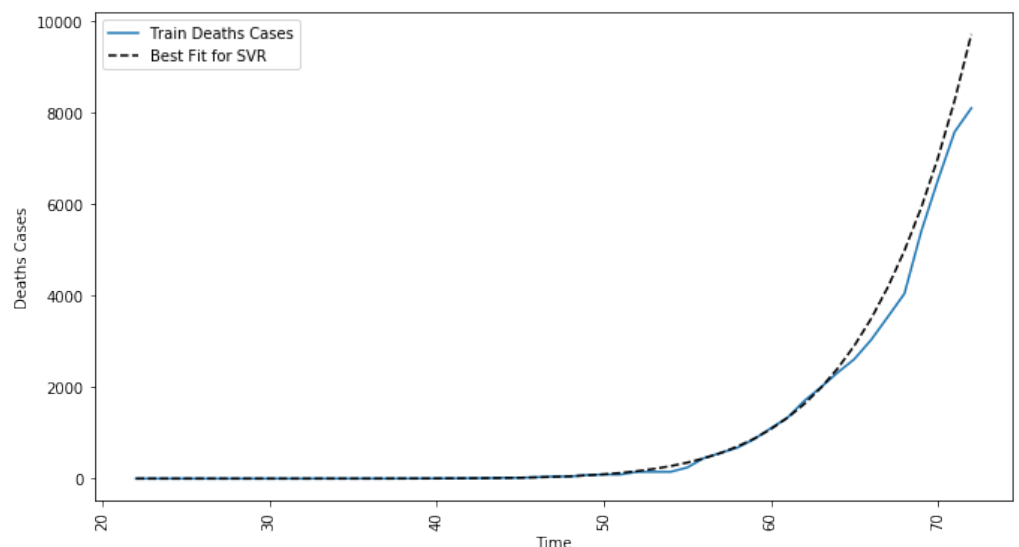


Figure 19- Résultat de la prédiction du modèle SVM pour les décès en France

6 Conclusions

Ce travail a été commencé durant les premiers jours du mois de mars, et depuis l'évolution de la pandémie a été vertigineuse.

Deux choses sont certaines : la première, les courbes des cas décédés et des cas des personnes contaminées sont exponentielles. L'autre que la mortalité du virus dépend strictement de mesures prises par les Etats de chaque pays.

Par rapport à nos modèles, on voit qu'avec les données mondiales on arrive à construire des modèles assez performant. Par contre, lorsqu'il s'agit de faire des modèles avec les données des pays individuels on a du mal à trouver. Certainement, la quantité de données récoltées n'est pas encore suffisante pour entraîner nos modèles de manière pertinente.

Pour l'instant, on est encore dans l'oeil de la tempête donc on n'a pas assez de recul pour en tirer des conclusions définitives. Reste à voir comment les choses vont évoluer, les analyser et apprendre des erreurs pour éviter qu'une nouvelle pandémie frappe le monde avec la puissance de celle-là.

7 Références

- What Is Time Series Forecasting? , Jason Brownlee
<https://machinelearningmastery.com/time-series-forecasting/>
- Time Series analysis tsa
<https://www.statsmodels.org/dev/tsa.html>
- SVM Support Vector Machine, Ricco Rakotomalala
<http://eric.univ-lyon2.fr/~ricco/cours/slides/svm.pdf>
- <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>
- COVID19 Global Forecasting (Week 2)
Forecast daily COVID-19 spread in regions around world
<https://www.kaggle.com/c/covid19-global-forecasting-week-2>

8 Annexe

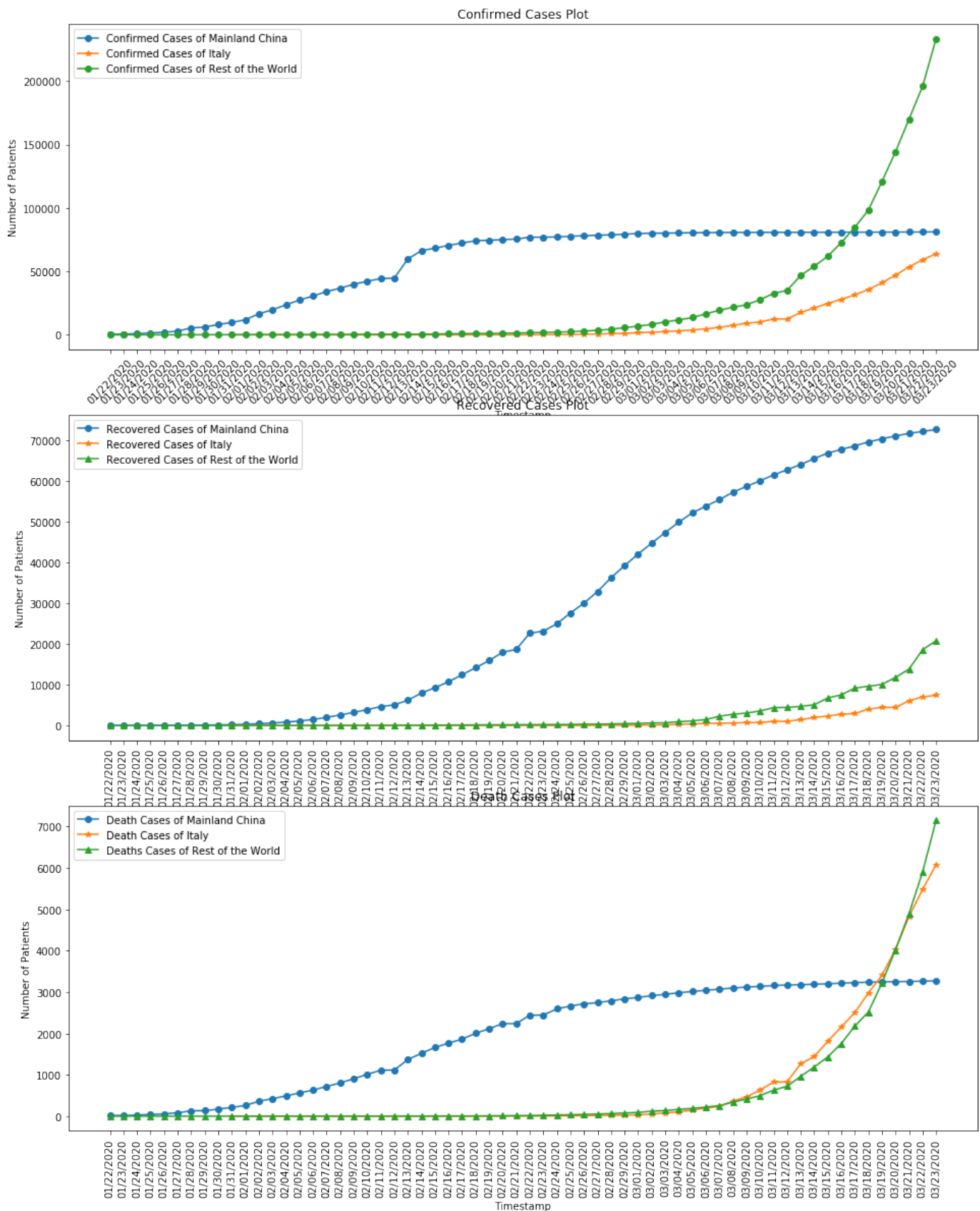


Figure 1- Comparaison entre la Chine , Italie et le reste du monde

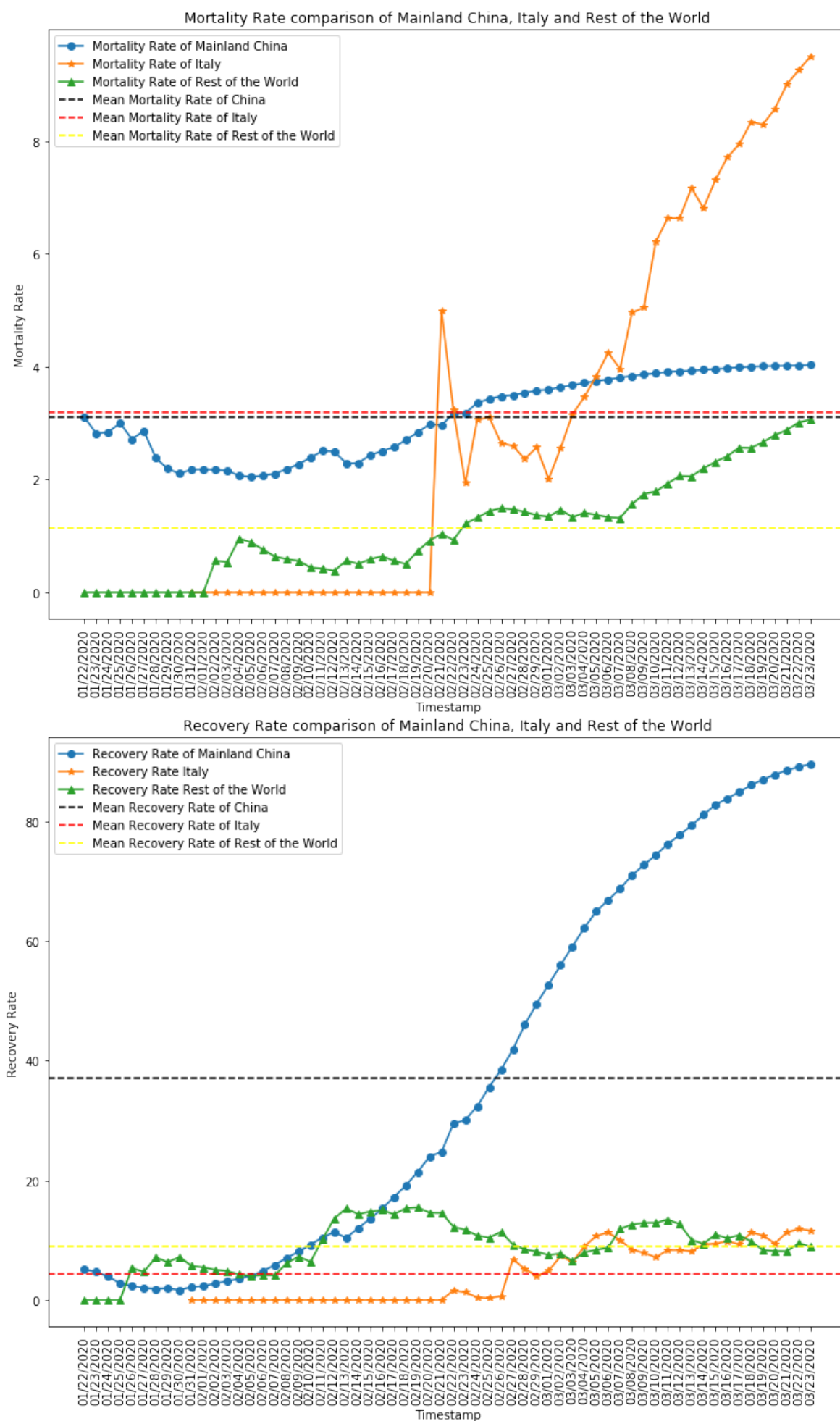
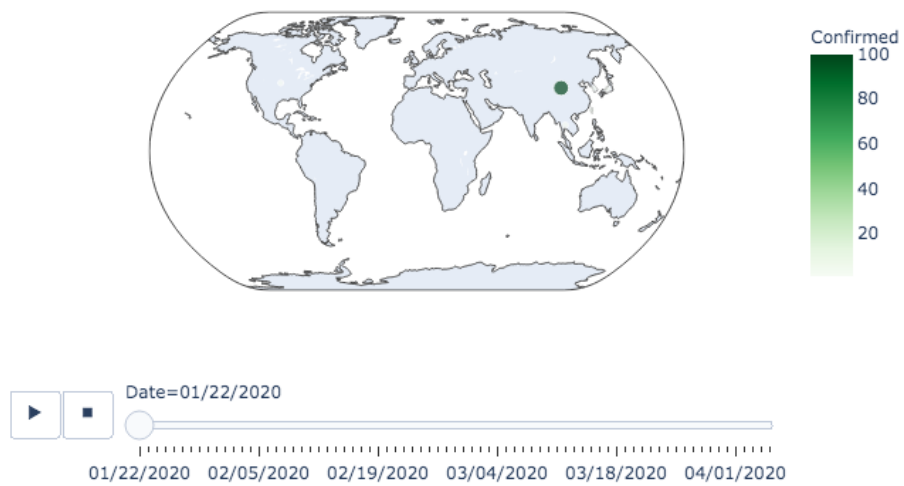
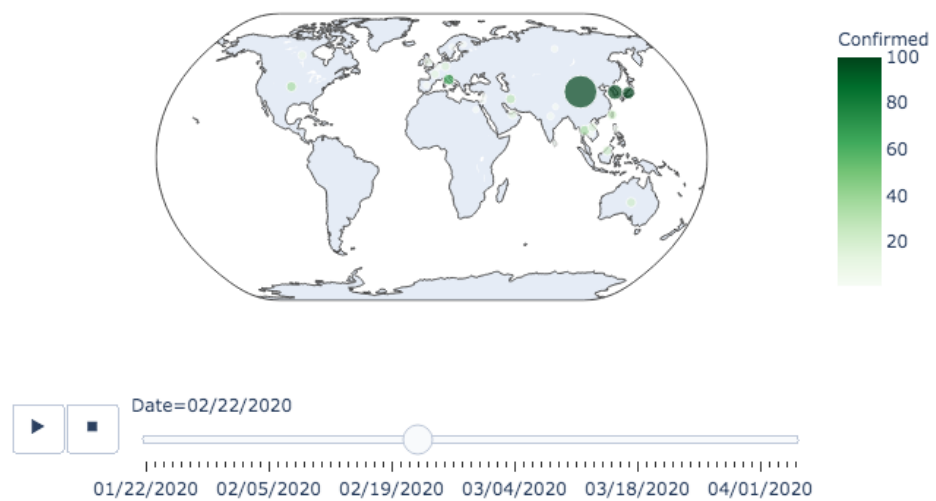


Figure 2- Comparaison entre la Chine, Italie et le reste du monde

COVID-19: Cases Over Time



COVID-19: Cases Over Time



COVID-19: Cases Over Time

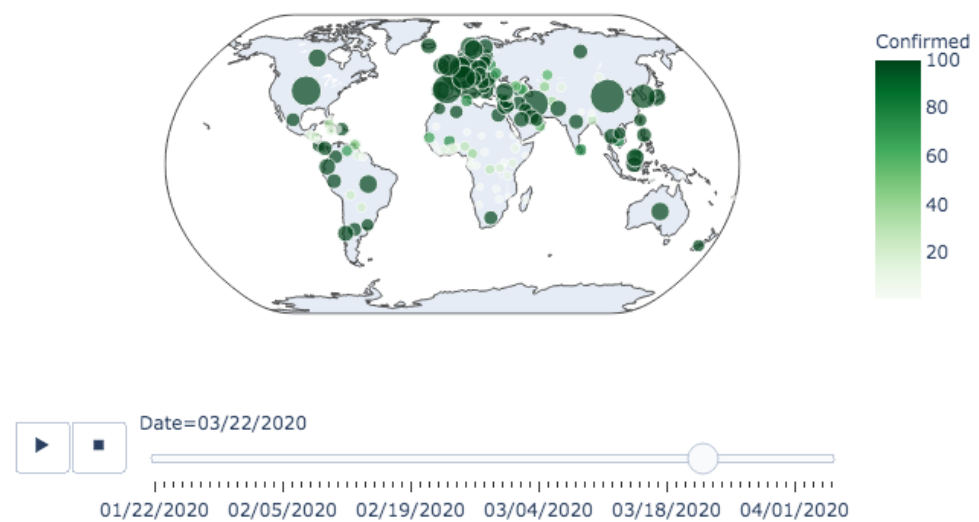


Figure 3- Ligne temporelle des cas confirmés de Covid_19 dans le monde