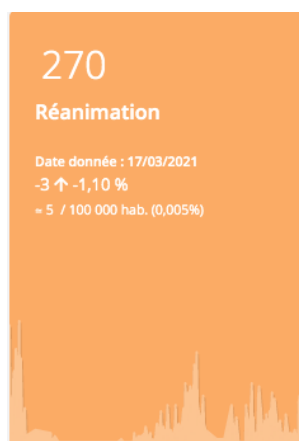
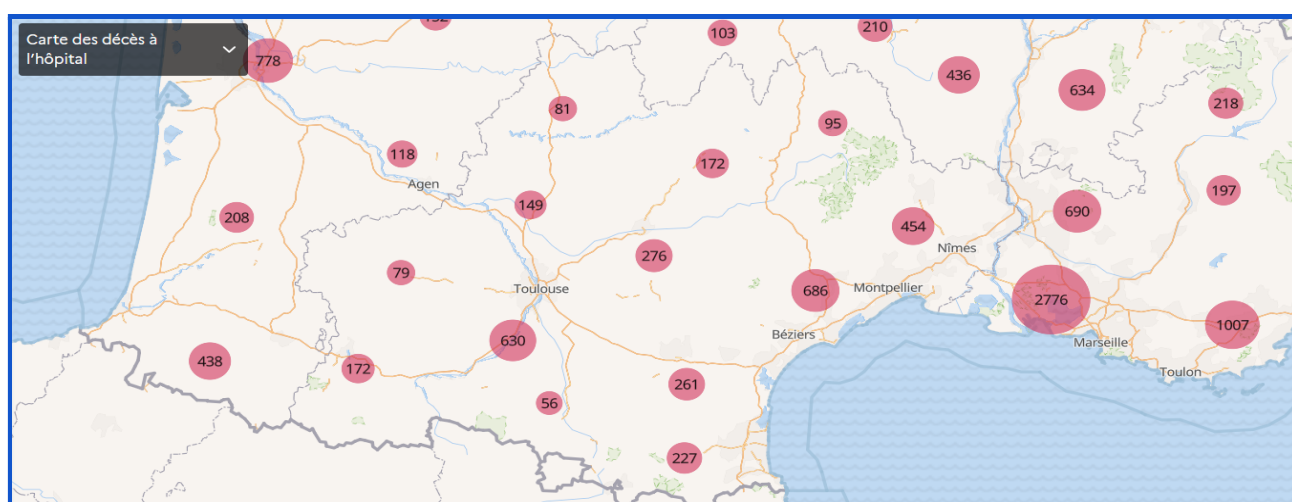


ANALYSE DE DONNÉES COVID-19 EN LA REGION DE L'OCCITANIE AVEC POWER BI

CASANOVA MARROQUIN Stephanya, RAMOS SEGURA YEIMI
{stephanya.casanova-marroquin, yeimi.ramos-segura}@univ-lyon2.fr



INTRODUCTION

COVID, SARS-CoV-2, agent pathogène qui donne l'origine d'une pneumonie atypique émergente, a été identifié en janvier 2020 en Chine .

Le 11 mars 2020, l'épidémie de Covid-19 est déclarée pandémie par l'OMS, et dans ce moment, l'OMS a demandé des mesures de protection essentielles pour prévenir la saturation des services de soins intensifs et renforcer l'hygiène préventive (suppression des contacts physiques, bises et poignées de mains, fin des attroupements et des grandes manifestations ainsi que des déplacements et voyages non indispensables, promotion du lavage des mains, mise en application de quarantaine, etc.).

COVID EN FRANCE

A la date 03/17/2021 en France, 25 314 personnes sont hospitalisées détectées positives au Covid-19, et 4 219 sont jugées dans un état critique. À ce jour, 91 565 personnes sont décédées.

COVID EN OCCITANIE

Occitanie, région de 13 départements, qui compte avec 6 009 231 habitants (population municipale au 1er janvier 2021), et constitue la cinquième région française (et métropolitaine) la plus peuplée.

Se trouve aujourd'hui avec (mars 17 2021) :

- 1 338 personnes hospitalisées détectées positives au Covid-19
- 270 sont jugées dans un état critique
- Après hospitalisation, 15 372 patients sont de retour à leur domicile
- À cette même date, 3 426 personnes sont décédées à l'hôpital depuis le début de l'épidémie

VACCINATION EN OCCITANIE

Depuis le début de la vaccination en France le 26 décembre 2020, 1 184 510 personnes ont ainsi été vaccinées. En Occitanie, 117 357 personnes ont déjà été vaccinées.

DATA

Les données utilisées pour la réalisation de ce projet ont été prises de la page d'OpenDatasoft. Opendatasoft est une société française proposant une solution Open Data complète : *publication - visualisation - partage - réutilisation des données*.

Les fichiers avec lesquels nous avons travaillé sont :

1. Données hospitalières relatives à l'épidémie de COVID-19 en France

Les données hospitalières quotidiennes relatives à l'épidémie du COVID-19 par département et sexe du patient : nombre de patients hospitalisés, nombre de personnes actuellement en réanimation ou soins intensifs, nombre cumulé de personnes retournées à domicile, nombre cumulé de personnes décédées.

Web page :

[https://public.opendatasoft.com/explore/dataset/donnees-hospitalieres-covid-19-dep-franc\[...\]
live.countrycode_iso_3166_1_alpha3&disjunctive.nom_dep_min](https://public.opendatasoft.com/explore/dataset/donnees-hospitalieres-covid-19-dep-franc[...]/live.countrycode_iso_3166_1_alpha3&disjunctive.nom_dep_min)

2. Indicateurs de suivi de l'épidémie de COVID-19 en France

Cette carte est construite sur la base de 4 indicateurs et est complétée par une analyse de risques. Ces quatre indicateurs permettent de mesurer l'intensité de l'épidémie sur le territoire et ainsi de décider d'un re confinement. Les données sont issues du Ministère des Solidarités et de la Santé.

Les indicateurs sont les suivants :

A. Activité épidémique (taux d'incidence)

Le taux d'incidence correspond au nombre de tests virologiques positifs pour 100.000 habitants sur une semaine. Il est calculé, à l'échelle départementale, de la manière suivante : $(100000 \times \text{nombre de cas positif}) / \text{population}$.

Trois niveaux ont été fixés pour cet indicateur :

vert : en dessous de 10 personnes testées positives sur 100.000 personnes testées, sur une semaine glissante ;

orange : au-delà de 10 personnes testées positives sur 100.000 personnes testées, sur une semaine glissante ;

rouge : au-delà de 50 personnes testées positives sur 100.000 personnes testées, sur une semaine glissante.

B. Taux de positivité des tests virologiques

Le taux de positivité correspond au nombre de personnes testées positives sur une semaine. Il est calculé, à l'échelle départementale, de la manière suivante : $(100 \times \text{nombre de tests positifs}) / \text{nombre de tests réalisés}$.

Trois niveaux ont été fixés pour cet indicateur :

vert : taux de positivité entre 0 et 5% ;

orange : taux de positivité entre 5 et 10% ;

rouge : taux de positivité supérieur à 10%.

C. Facteur de reproduction du virus (évolution du R0)

Nombre de personnes contaminées par chaque malade. Cet indicateur est calculé à l'échelle régionale.

Trois niveaux ont été fixés pour cet indicateur :

Vert : R0 entre 0 et 1 ;

Orange : R0 entre 1 et 1,5 ;

Rouge : R0 supérieur à 1,5.

D. Tension hospitalière sur la capacité en réanimation

Taux d'occupation des lits en réanimation/SI/SC par des patients COVID par rapport à la capacité initiale en réanimation/SI/SC. Cet indicateur est calculé à l'échelle régionale.

Trois niveaux ont été fixés pour cet indicateur :

Vert : taux d'occupation compris entre 0 et 40% ;

Orange : taux d'occupation compris entre 40 et 60% ;

Rouge : taux d'occupation supérieur à 60%.

Web page :

https://public.opendatasoft.com/explore/dataset/coronavirus-tranche-age-urgences-sosmedec-ins-dep-france/table/?disjunctive.nom_dep_min

3. Données relatives aux vaccinations par région - COVID-19 - France

Le jeu de données contient le nombre cumulé de personnes vaccinées à la maille régionale. Le nombre de vaccinations en France est obtenu par recueil quotidien auprès des ARS des données régionales de vaccination, recensées auprès des établissements pour personnes âgées (vaccination des résidents et des professionnels des établissements) et des centres de vaccination (vaccination des professionnels de santé).

Webpage:[https://pubic.opendatasoft.com/explore/dataset/covid-19-france-vaccinations-region/tabl\[...\]SwiZGlzcGxheUxlZ2VuZCI6dHJ1ZSwiYWxpZ25Nb250aCI6dHJ1ZX0%3D](https://pubic.opendatasoft.com/explore/dataset/covid-19-france-vaccinations-region/tabl[...]SwiZGlzcGxheUxlZ2VuZCI6dHJ1ZSwiYWxpZ25Nb250aCI6dHJ1ZX0%3D)

4. Vacsi-v-reg-2021-03-10-17h25

Le jeu de données issues du système d'information Vaccin Covid permettent de dénombrer en temps quasi réel (J-1), le nombre de personnes ayant reçu une injection de vaccin anti-covid en tenant compte du nombre de doses reçues, du vaccin ainsi que du niveau géographique (national, régional et départemental).

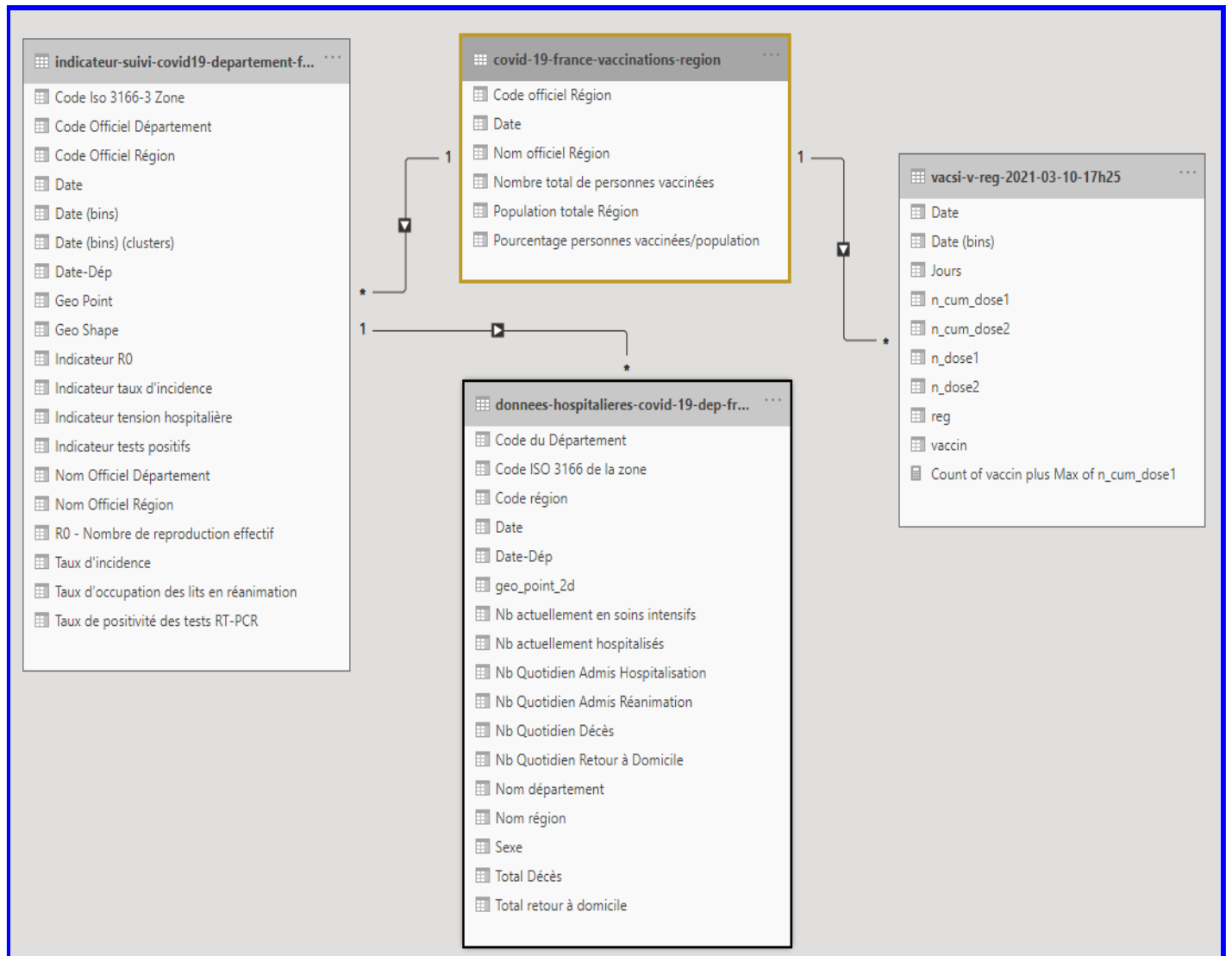
Les vaccins sont codifiés de la façon suivante :

- 0 : Tous vaccins
- 1 : COMIRNATY Pfizer/BioNTech
- 2 : Moderna
- 3 : AstraZeneka

Web page : <https://www.data.gouv.fr/fr/datasets/donnees-relatives-aux-personnes-vaccinees-contre-la-covid-19-1/>

MODÈLE DE DONNÉES

Dans la figure suivante, nous présentons le modèle de données du projet. En cela, nous pouvons trouver les 4 tables avec lesquelles nous travaillons et leurs relations.



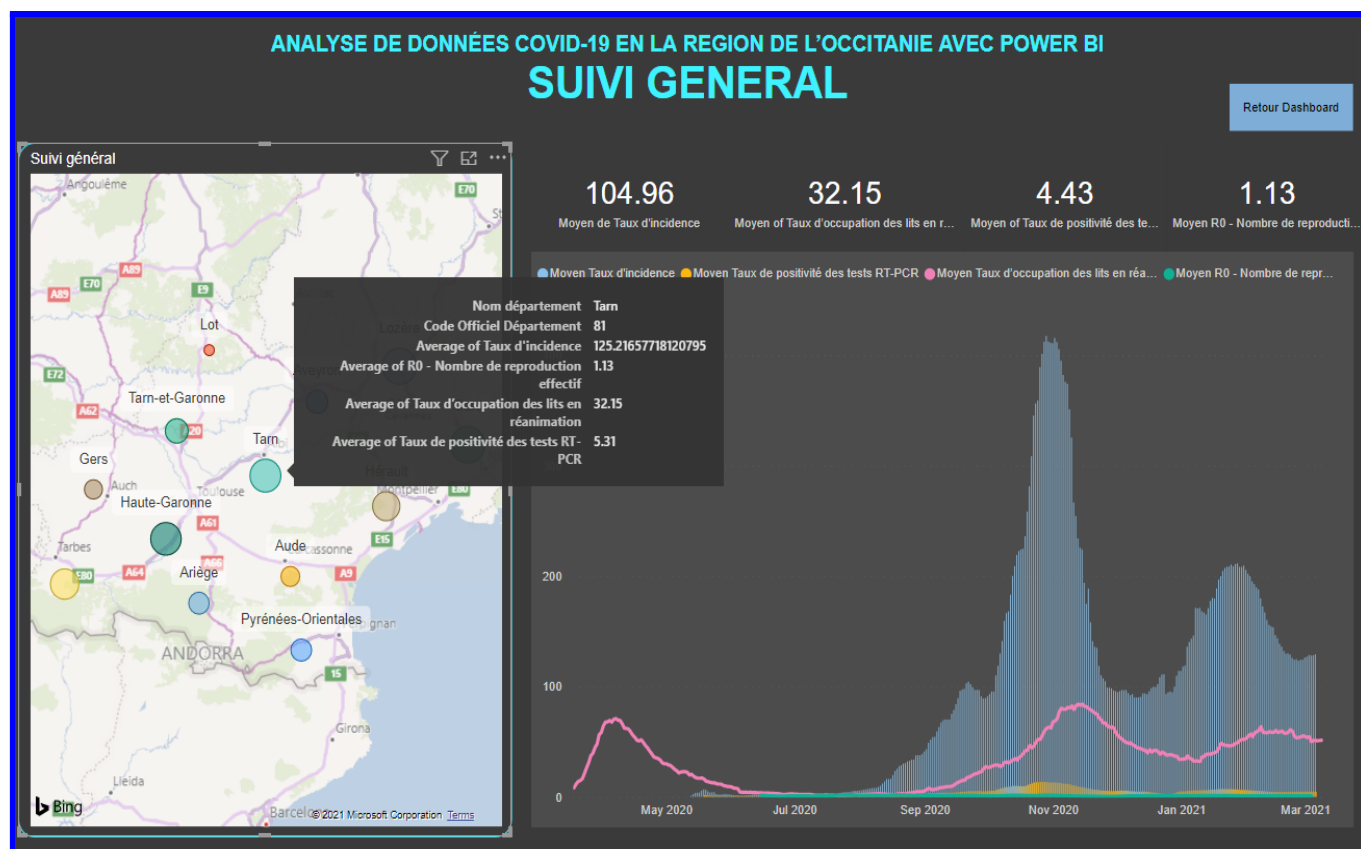
PAGES DU REPORT

Pour le projet nous avons créé les suivants pages :

- **Suivi Général,**
- **Gestion Hospitalière,**
- **Vaccination,**
- **Clustering Gestion Emergence,**
- **Prédiction Vaccination Total et**
- **Prédiction R**

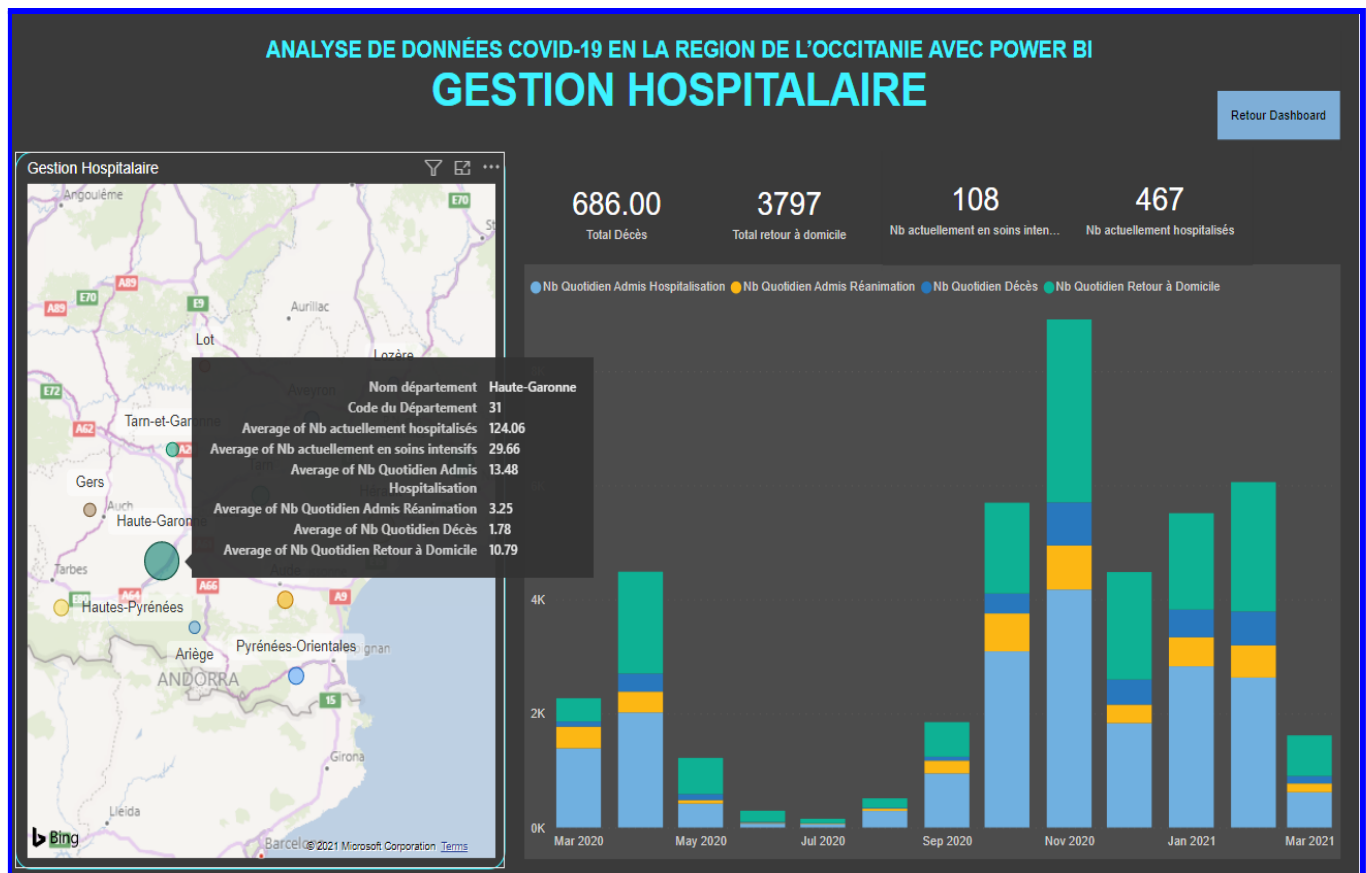
La page suivante concerne le fichier "indicateur-suivi-covid19-departemen-france.csv". Cette page présente 4 indicateurs : *Moyen de taux d'incidence dans la région, moyen de taux d'occupation de lits, moyen de test de positivité de tests, moyen d'indice de reproduction du virus.*

Les 2 figures montrent : L'historique moyen des indicateurs déjà mentionnés et la carte géographique permet d'observer le valeur moyen de chaque indicateur dans chaque département.



La page suivante concerne le fichier “donnees-hospitalaires-covid-19-dep-france.csv”. Cette page présente 4 indicateurs : total décès, total retour au domicile, Nombre de patients en soins intensifs et nombre d'hospitalisations à la date.

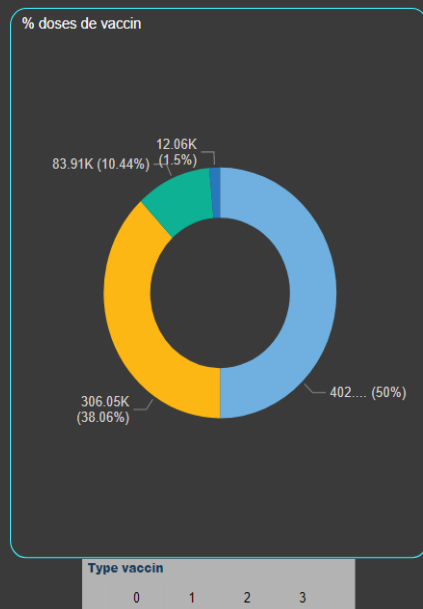
Les 2 figures montrent: L'historique moyen des indicateurs déjà mentionnés et la carte géographique permet d'observer la valeur moyenne de chaque indicateur dans chaque département.



La page suivante concerne le fichier “vacsi-v-reg-2021-03-10-17h25.csv”. Cette page présente 4 indicateurs : Population Total de la région d’Occitanie, Numéro de doses appliquées jusqu'au 10 mars, les différents types de vaccins utilisés et le nombre de jours qui ont passé depuis que la vaccination a commencé dans la région.

Les 2 figures montrent : % de chaque vaccin utilisé et l'historique de vaccination dans la région pour chaque type de vaccin.

ANALYSE DE DONNÉES COVID-19 EN LA REGION DE L'OCCITANIE AVEC POWER BI VACCINATION



6.01M

Population totale Région

804K

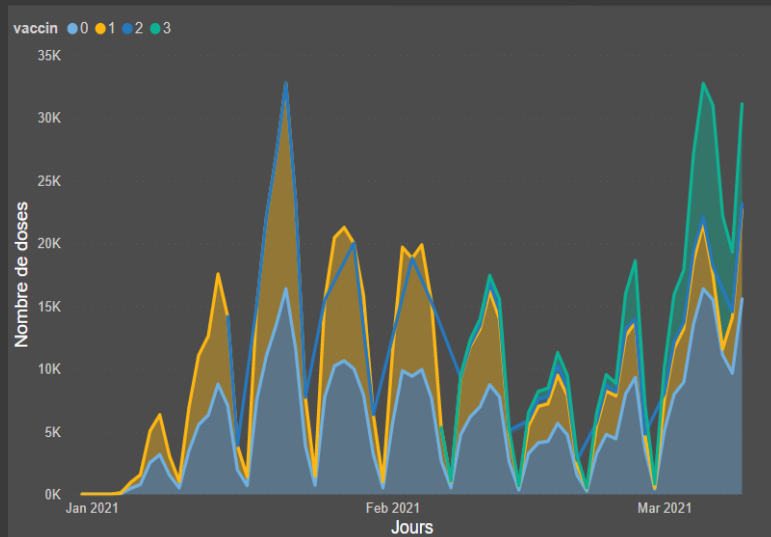
Nombre total 1re doses appliqués

4

Nombre de vaccins

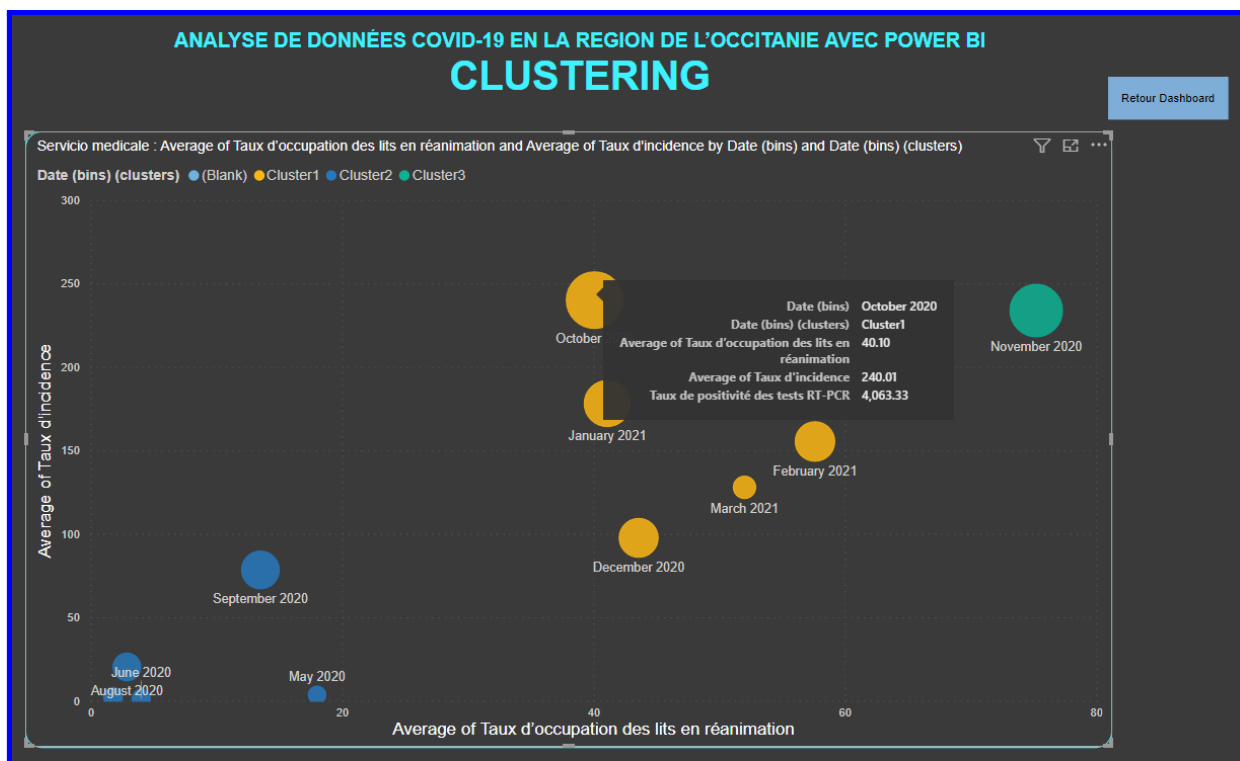
69

Jours



CLUSTERING GESTION HOSPITALIÈRE

La page suivante concerne le clustering fait sur les données de “donnees-hospitalaires-covid-19-dep-france.csv”. Dans la graphique, nous pouvons observer 3 clusters qui correspondent à la gestion médicale entre mai 2020 et mars 2021. Nous pouvons observer que le cluster 3 correspond à un 2^{ème} pic de l'émergence en novembre de 2020, le cluster 1 correspond à la période de mitigation de l'émergence générée dans le 2^{ème} pic, et le cluster 2 correspond à autre période de mitigation après de 1^{ère} pic.



PREVISIÓN DE VACCINATION TOTAL EN OCCITANIE

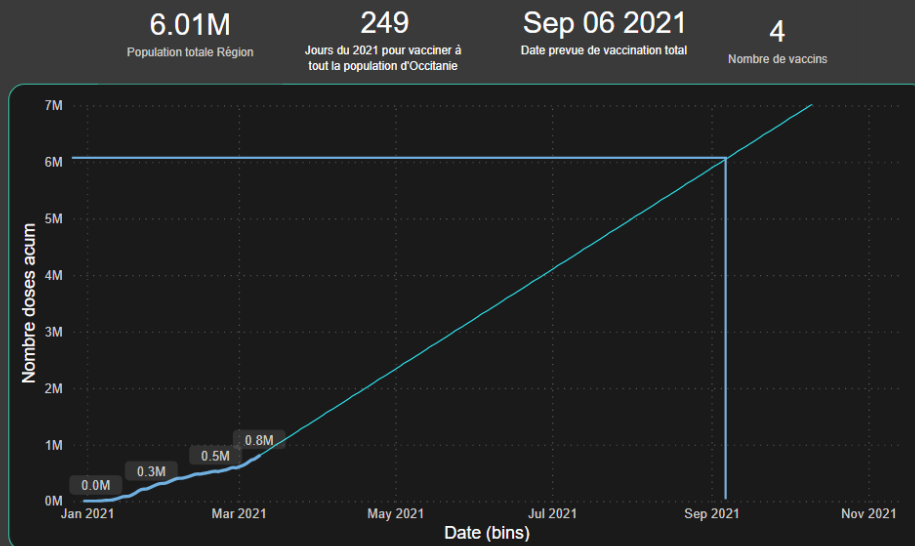
Cette page concerne la prédiction faite à partir du fichier “vacsi-v-reg-2021-03-10-17h25.csv”. Cette prédiction cherche à estimer la date à laquelle la population de la région occitanie sera complètement vaccinée. Pour obtenir cette prédiction, nous avons utilisé le composants “Analytics” et la fonction de “forecasting”. Nous avons fait une prédiction pour 8 mois et nous avons défini un intervalle de confiance de 99%.

Le résultats obtenus sont :

Numéro de jours pour vacciner tout la population : 249

Date de fin de vaccination : sep 06 de 2021

ANALYSE DE DONNÉES COVID-19 EN LA REGION DE L'OCCITANIE AVEC POWER BI PREDICTION



PREDICTION AVEC R

Dans cette page, nous avons essayé de faire la même prédiction de la dernière page mais cette fois en utilisant la librairie “forecast” de R. A continuation le code utilisée :

```
# The following code to create a dataframe and remove duplicated rows is  
always executed and acts as a preamble for your script:
```

```
# dataset <- data.frame(Jours, n_cum_dose1)  
# dataset <- unique(dataset)
```

```
# Paste or type your script code here:
```

```
library(forecast)
```

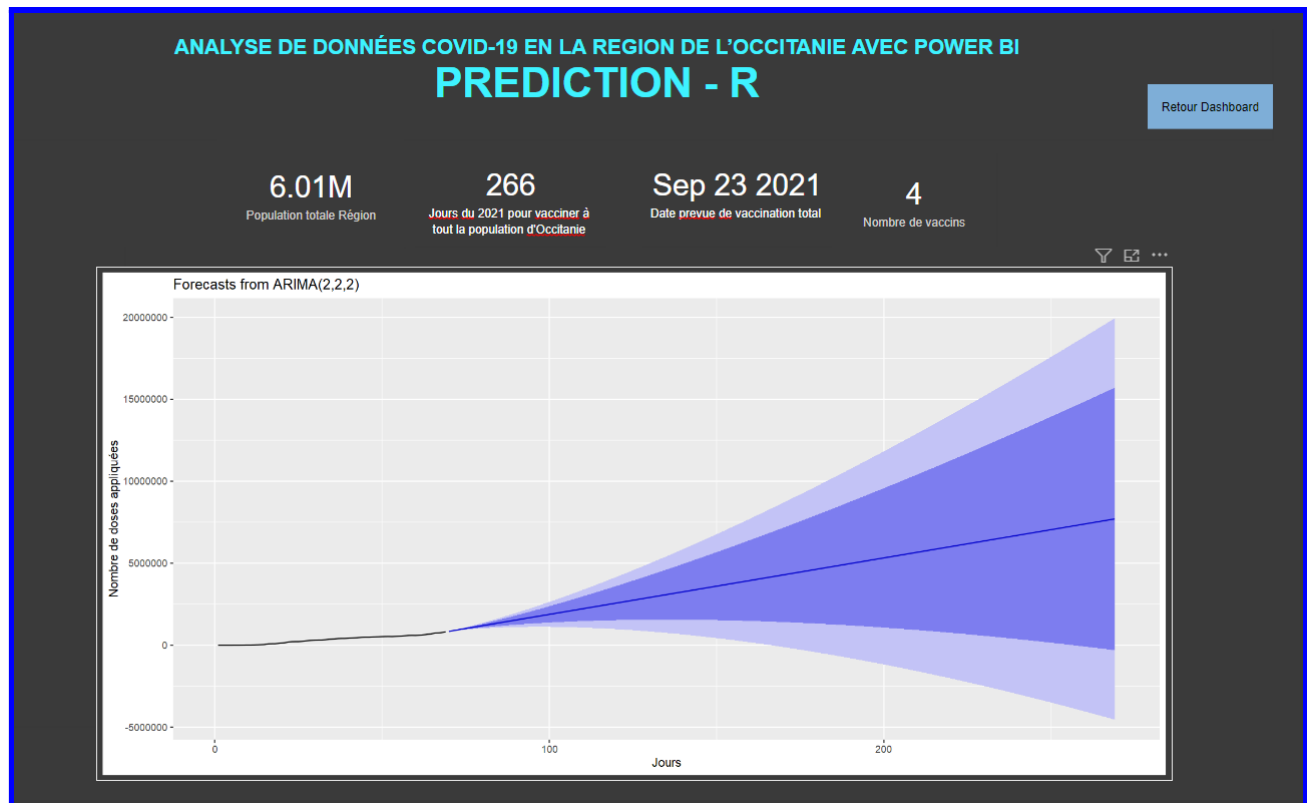
```
dataset <- aggregate( n_cum_dose1 ~ Jours, data = dataset, FUN = sum)  
ts <- ts(dataset$n_cum_dose1)  
arima.ts <- auto.arima(ts)
```

```
pred <- forecast(arima.ts, h = 200)
autoplot(pred,xlab = "Jours",ylab = "Nombre de doses appliquées")
```

La graphique suivante montre les résultats obtenus :

Numéro de jours pour vacciner tout la population : 266 contre 249 de la dernière page

Date de fin de vaccination : sep 23 de 2021 à différence du sep 06 de 2021 de la dernière page



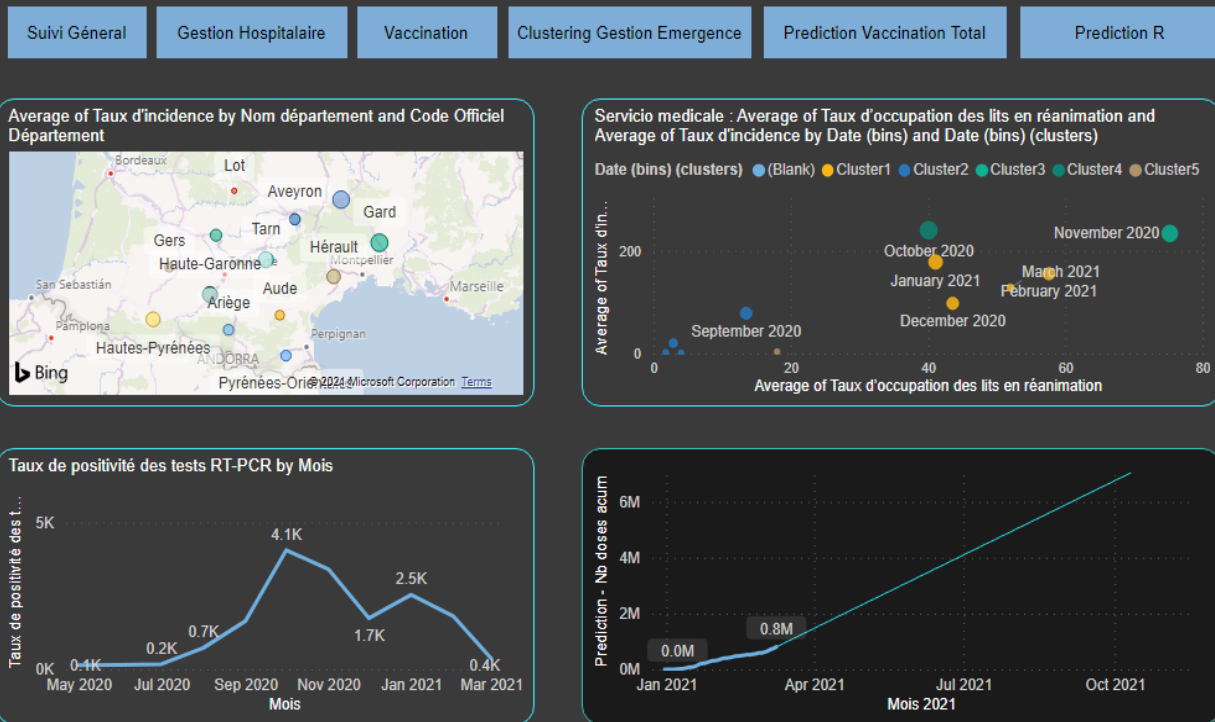
DASHBOARD COMPLETE

Dans le dashboard, nous trouvons :

- Taux d'incidence par département dans l'Occitanie
- Taux de positivité des test par mois depuis de may 2020
- Clustering pour observer la gestion de service médicale en considérant le taux d'occupation de lits et le taux d'incidence par moi
- Prédiction de jours et date final de la vaccination total de l'Occitanie

- 6 boutons qui permettent d'accéder à chacune des pages : **Suivi Général, Gestion Hospitalière, Vaccination, Clustering Gestion Emergence, Prédiction Vaccination Total et Prédiction R.**

ANALYSE DE DONNÉES COVID-19 EN LA REGION DE L'OCCITANIE AVEC POWER BI



REFERENCES

<https://www.gouvernement.fr/info-coronavirus/carte-et-donnees>

<https://data.laregion.fr/pages/dashboard-covid/>

<https://www.youtube.com/watch?v=8uQ8CbScyrM>

<https://github.com/Pitsillides91/Python-Tutorials/tree/master/CoronaVirus>

<https://www.data.gouv.fr/fr/datasets/donnees-relatives-aux-personnes-vaccinees-contre-la-covid-19>

-1/