



Projet Framework Big Data

Sur le thème :

Analyse des Tweets des PROVAX et des ANTIVAX Covid-19

Réalisé Par :

Lydia Ait-Abdelkader

Yanis Hammoudi

Année Universitaire : 2021-2022

Table des matières

Introduction	3
Architecture	3
Ingestion de données	5
Requêtes Hive.....	8
Conclusion	14

Introduction

Depuis début 2021 et l'arrivée du vaccin contre le COVID-19, un débat intense anime les réseaux sociaux, notamment twitter, où on a l'impression que 2 camps se sont formés. D'un côté ceux qui sont pour le vaccin, d'un autre ceux qui sont contre le vaccin. Des manifestations partout dans le monde et notamment en France ont souvent lieu pour exprimer leur opposition au pass « Vaccinal ». Dans ce projet, nous voulons résoudre certains points qui sont flous, à travers l'analyse des tweets, nous allons voir les raisons qui poussent les français à se faire vacciner ainsi que celles de ceux qui s'y opposent. Nous analyserons donc tout cela en détail dans ce qui suit.

Architecture

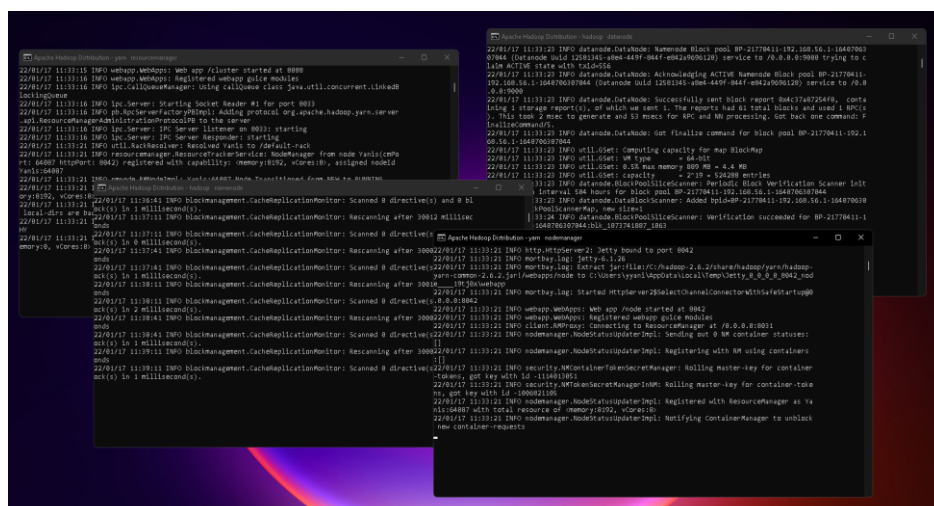
Nous avons choisi de réaliser notre projet en local from scratch. Nous avons donc installé chaque brique séparément sur l'ordinateur.

Nous avons en premier lieu installé HADOOP, les étapes que nous avons suivies sont :

- Installation de Java 8 et configuration des variables d'environnement.
- Installation d'éclipse et du plugin associé (hadoop2x-eclipse-plugin-master).
- Installation d'apache.
- Installation d'HADOOP version 2.6 et configuration des fichiers core-site.xml, hdfs-site.xml, mapred.xml et yarn-site.xml.

```
C:\>hadoop version
Hadoop 2.6.2
Subversion https://git-wip-us.apache.org/repos/asf/hadoop.git -r 0cfd050febe4a30b1ee1551dcc527589509fb681
Compiled by jenkins on 2015-10-22T00:42Z
Compiled with protoc 2.5.0
From source with checksum f9ebb94bf5bf9bec892825ede28bac
This command was run using /C:/hadoop-2.6.2/share/hadoop/common/hadoop-common-2.6.2.jar
```

- Lancement d'HADOOP en tapant les commandes start-dfs et start-yarn.



Après cela, nous avons installé Hive qui est une infrastructure d'entrepôt de données permettant de faire des requêtes via un langage proche syntaxiquement de SQL. Les étapes d'installation que nous avons suivies sont comme suit :

- Installation d'Apache derby et extraction des fichiers à mettre dans le même répertoire où se situe Hadoop.
- Téléchargement d'Apache Hive depuis le site officiel et extraction de l'archive dans le répertoire où se situe Hadoop.
- Définition des variables d'environnement HIVE_HOME, DERBY_HOME, HIVE_LIB, HIVE_BIN ainsi que la variable système HADOOP_USER_CLASSPATH_FIRST : « vrai ».
- Configuration de Hive en allant dans le répertoire des bibliothèques Derby et en copiant tous les fichiers .jar pour ensuite les coller dans le répertoire des bibliothèques Hive.
- Configuration de hive-site.xml.
- Démarrage du serveur derby avec la commande C:\derby\bin\StartNetworkServer -h 0.0.0.0.
- Verification de la version de hive.

```
Beeline version 2.1.0 by Apache Hive
```

Enfin, nous avons installé Flume qui est une solution de collecte, agrégation et transfert de gros volumes de données. Il a été conçu pour gérer des débits importants avec une fonctionnalité native d'écriture dans HDFS. Pour gérer ces gros volumes/débits, il se doit d'être très scalable, et donc distribué. L'outil fait partie de l'écosystème Big Data open source Hadoop.

Les étapes d'installation de Flume sont les suivantes :

- Télécharger Apache Flume depuis le site officiel <http://flume.apache.org/download.html> et le mettre dans le répertoire C:\.
- Définir des variables d'environnement FLUME_HOME, FLUME_CONF, PATH.
- Aller dans le dossier C:\apache-flume-1.9.0-bin\conf, ouvrir le fichier log4j.properties et faire les changements suivant :
flume.root.logger=DEBUG,console
#flume.root.logger=INFO,LOGFILE
- Renommer les fichiers flume-conf.properties.template en flume-conf.properties et flume-env.ps1.template en flume-env.ps1
- Flume est maintenant installé, nous vérifions sa version avec la commande 'flume-ng version'

```

C:\apache-flume-1.9.0-bin\bin>flume-ng version

C:\apache-flume-1.9.0-bin\bin>powershell.exe -NoProfile -InputFormat none -ExecutionPolicy
flume-1.9.0-bin\bin\flume-ng.ps1 version

WARN: Config directory not set. Defaulting to C:\apache-flume-1.9.0-bin\conf
Sourcing environment configuration script C:\apache-flume-1.9.0-bin\conf\flume-env.ps1

Running FLUME version :
  class: org.apache.flume.tools.VersionInfo
  arguments:

Flume 1.9.0

```

Ingestion de données

Nous avons récupéré les données depuis twitter avec Apache Flume, en faisant deux requêtes, la première pour récupérer les tweets ANTIVAX et la seconde pour les tweets PROVAX.

Afin d'obtenir les tweets, nous avons créé une application Twitter depuis le site <https://developer.twitter.com/>. Les étapes sont énumérées ci-dessous :

- Nous avons cliqué sur Create New APP. Nous avons été redirigés vers une fenêtre où nous avons renseigné les coordonnées de l'application via un formulaire.
- L'application a bien été créée avec les détails indiqués.
- En dessous de keys and Access Tokens onglet en bas de page, nous avons cliqué sur le bouton nommé Create my access token qui a généré un jeton d'accès.
- Cela a permis d'ouvrir une page contenant consumerKey , consumerSecret , accessToken et accessTokenSecret.

En premier lieu, nous avons créé deux répertoires hdfs dans lesquels les données extraites depuis twitter seront sauvegardées, un répertoire pour les tweets ProVax et un autre pour les ANTIVAX en utilisant les commandes :

```
hdfs dfs -mkdir hdfs://localhost:50070/user/Hadoop/twitter_data/ProVax.
```

```
hdfs dfs -mkdir hdfs://localhost:50070/user/Hadoop/twitter_data/AntiVax.
```

Ensuite nous avons modifié, l'agent flume , le fichier de configuration 'twitter.conf' dans le dossier conf de Flume :

- ProVax :

```
# Naming the components on the current agent.
TwitterAgent.sources = Twitter
TwitterAgent.channels = MemChannel
TwitterAgent.sinks = HDFS

# Describing/Configuring the source
TwitterAgent.sources.Twitter.type = org.apache.flume.source.twitter.TwitterSource
TwitterAgent.sources.Twitter.channels = MemChannel
TwitterAgent.sources.Twitter.consumerKey = CwmlH5Gp3tjy8H4wXllyz3D9o
TwitterAgent.sources.Twitter.consumerSecret = 42VVK4BzN5T4gXNWN4oXN67KjWd73o07k411eTeechNkH20P
TwitterAgent.sources.Twitter.accessToken = 1475010F02350276074_6130M4M0U4M0p3p8C0010000m12W
TwitterAgent.sources.Twitter.accessTokenSecret = K2gXNMF1W2Q6fyj4G6N4u8J4N1L40Ug10U5500U107
TwitterAgent.sources.Twitter.language = fr
TwitterAgent.sources.Twitter.keywords = vaccin , covid , je , suis , pour , pass

# Describing/Configuring the sink

TwitterAgent.sinks.HDFS.type = hdfs
TwitterAgent.sinks.HDFS.hdfs.path = hdfs://localhost:50070/user/Hadoop/twitter_data/ProVax

TwitterAgent.sinks.HDFS.hdfs.fileType = DataStream
TwitterAgent.sinks.HDFS.hdfs.writeFormat = Text
TwitterAgent.sinks.HDFS.hdfs.batchSize = 100
TwitterAgent.sinks.HDFS.hdfs.rollSize = 0
TwitterAgent.sinks.HDFS.hdfs.rollCount = 100

# Describing/Configuring the channel
TwitterAgent.channels.MemChannel.type = memory
TwitterAgent.channels.MemChannel.capacity = 10000
TwitterAgent.channels.MemChannel.transactionCapacity = 1000

# Binding the source and sink to the channel
TwitterAgent.sources.Twitter.channels = MemChannel
TwitterAgent.sinks.HDFS.channel = MemChannel
```

- AntiVax :

```
# Naming the components on the current agent.
TwitterAgent.sources = Twitter
TwitterAgent.channels = MemChannel
TwitterAgent.sinks = HDFS

# Describing/Configuring the source
TwitterAgent.sources.Twitter.type = org.apache.flume.source.twitter.TwitterSource
TwitterAgent.sources.Twitter.channels = MemChannel
TwitterAgent.sources.Twitter.consumerKey = CwmlH5Gp3tjy8H4wXllyz3D9o
TwitterAgent.sources.Twitter.consumerSecret = 42VVK4BzN5T4gXNWN4oXN67KjWd73o07k411eTeechNkH20P/
TwitterAgent.sources.Twitter.accessToken = 1475010F02350276074_6130M4M0U4M0p3p8C0010000m12W
TwitterAgent.sources.Twitter.accessTokenSecret = K2gXNMF1W2Q6fyj4G6N4u8J4N1L40Ug10U5500U107
TwitterAgent.sources.Twitter.language = fr
TwitterAgent.sources.Twitter.keywords = vaccin , covid , je , suis , contre , anti, pass

# Describing/Configuring the sink

TwitterAgent.sinks.HDFS.type = hdfs
TwitterAgent.sinks.HDFS.hdfs.path = hdfs://localhost:50070/user/Hadoop/twitter_data/AntiVax

TwitterAgent.sinks.HDFS.hdfs.fileType = DataStream
TwitterAgent.sinks.HDFS.hdfs.writeFormat = Text
TwitterAgent.sinks.HDFS.hdfs.batchSize = 100
TwitterAgent.sinks.HDFS.hdfs.rollSize = 0
TwitterAgent.sinks.HDFS.hdfs.rollCount = 100

# Describing/Configuring the channel
TwitterAgent.channels.MemChannel.type = memory
TwitterAgent.channels.MemChannel.capacity = 10000
TwitterAgent.channels.MemChannel.transactionCapacity = 1000

# Binding the source and sink to the channel
TwitterAgent.sources.Twitter.channels = MemChannel
TwitterAgent.sinks.HDFS.channel = MemChannel
```

Après avoir configuré l'agent Flume , nous allons lancé le script Flume pour générer les données depuis twitter avec la commande suivante :

flume-ng agent -n TwitterAgent -c conf -f C:/apache-flume-1.9.0-bin/conf/twitter.conf -property flume.root.logger=DEBUG,console

```
C:\apache-flume-1.9.0-bin>flume-ng agent -n TwitterAgent -c conf -f C:/apache-flume-1.9.0-bin/conf/twitter.conf -property flume.root.logger=DEBUG,console
```

Les données sont alors récupérées et stockées dans HDFS :

```
r.java:75]] Received:{"extended_tweet":{"extended_entities":{"media":[{"display_url":"pic.twitter.com/1l1fBhMB3p","indices":
es":[142,165],"sizes":{"small":{"w":680,"h":383,"resize":"fit"},"large":{"w":1920,"h":1080,"resize":"fit"},"thumb":{"w":
150,"h":150,"resize":"crop"},"medium":{"w":1200,"h":675,"resize":"fit"},"id_str":"1483070992646049800","expanded_url":"
https://twitter.com/sign112358stray/status/1483070997817597953/photo/1","media_url_https":"https://pbs.twimg.com/media/F
JTtNclakAgTnow.jpg","id":"1483070992646049800","type":"photo","media_url":"http://pbs.twimg.com/media/FJTtNclakAgTnow.jpg"
,"url":"https://t.co/1l1fBhMB3p"}]},"entities":{"urls":[],"hashtags":[],"media":[{"display_url":"pic.twitter.com/1l1fBhMB
B3p","indices":[142,165],"sizes":{"small":{"w":680,"h":383,"resize":"fit"},"large":{"w":1920,"h":1080,"resize":"fit"},"t
thumb":{"w":150,"h":150,"resize":"crop"},"medium":{"w":1200,"h":675,"resize":"fit"},"id_str":"1483070992646049800","expa
nded_url":"https://twitter.com/sign112358stray/status/1483070997817597953/photo/1","media_url_https":"https://pbs.twimg.
com/media/FJTtNclakAgTnow.jpg","id":"1483070992646049800","type":"photo","media_url":"http://pbs.twimg.com/media/FJTtNclak
AgTnow.jpg","url":"https://t.co/1l1fBhMB3p"}]},"user_mentions":[],"symbols":[],"full_text":"??????77?\\n1?????????????
?????????\\n?????OP?????????????\\n?????????????\\n?????????????\\n?????????????\\n?????????????\\n?????????????
?????\\u2026? https://t.co/1l1fBhMB3p","display_text_range":[0,141],"in_reply_to_status_id_str":null,"in_reply_to_status
_id":null,"created_at":"Mon Jan 17 13:37:34 +0000 2022","in_reply_to_user_id_str":null,"source":"<a href='\"http://twitte
r.com/download/iphone\"' rel='\"nofollow\"'>Twitter for iPhone</a>","retweeted":false,"geo":null,"filter
_level":"low","in_reply_to_screen_name":null,"is_quote_status":false,"id_str":"1483070997817597953","in_reply_to_user_id
":null,"favorite_count":0,"id":"1483070997817597953","text":"??????77?\\n1?????????????????????????????\\n?????OP????????????
?????????\\n?????????????\\u2026?????????????????????\\n?????????????????\\u2026 https://t.co/Ex7MZMCifv","place":null,"lang":"ja",
"quote_count":0,"favorited":false,"possibly_sensitive":false,"coordinates":null,"truncated":true,"timestamp_ms":"1642426
654663","reply_count":0,"entities":{"urls":[{"display_url":"twitter.com/i/web/status/1\\u2026","indices":[117,140],"expa
nded_url":"https://twitter.com/i/web/status/1483070997817597953","url":"https://t.co/Ex7MZMCifv"}],"hashtags":[],"user_me
ntions":[],"symbols":[],"display_text_range":[0,140],"contributors":null,"user":{"utc_offset":null,"friends_count":1581
,"profile_image_url_https":"https://pbs.twimg.com/profile_images/1475807396505485312/lxa2NmjV_normal.jpg","listed_count"
:9,"profile_background_image_url":"","default_profile_image":false,"favourites_count":6575,"description":"???-???-???-??
???????/Twitter????????????-???-????????????/?????????\\u2026?????????\\u2026","created_at":"Sun Sep 26 16:29:45 +0000 202
1","is_translator":false,"withheld_in_countries":[],"profile_background_image_url_https":"","protected":false,"screen_na
me":"sign112358stray","id_str":"1442164387700117520","profile_link_color":"1DA1F2","translator_type":"none","id":"1442164
387700117520","geo_enabled":false,"profile_background_color":"F5F8FA","lang":null,"profile_sidebar_border_color":"C0DEED",
"profile_text_color":"333333","verified":false,"profile_image_url":"http://pbs.twimg.com/profile_images/147580739650548
5312/lxa2NmjV_normal.jpg","time_zone":null,"url":null,"contributors_enabled":false,"profile_background_tile":false,"prof
```

On peut les apercevoir dans le localhost (<http://localhost:50070/>)

Hadoop Overview Datanodes Snapshot Startup Progress Utilities							
Browse Directory							
/user/hadoop/twitter_data/ProVax							Go!
Permission	Owner	Group	Size	Replication	Block Size	Name	
-rwxr-xr-x	Yanis	supergroup	668.45 KB	1	128 MB	FlumeData.1641038549908	
-rwxr-xr-x	Yanis	supergroup	667.8 KB	1	128 MB	FlumeData.1641038611097	
-rwxr-xr-x	Yanis	supergroup	673.21 KB	1	128 MB	FlumeData.1641038641998	
-rwxr-xr-x	Yanis	supergroup	708.03 KB	1	128 MB	FlumeData.1641038672991	
-rwxr-xr-x	Yanis	supergroup	682.57 KB	1	128 MB	FlumeData.1641038703937	
-rwxr-xr-x	Yanis	supergroup	699.31 KB	1	128 MB	FlumeData.1641038734937	
-rwxr-xr-x	Yanis	supergroup	649.46 KB	1	128 MB	FlumeData.1641038765908	
-rwxr-xr-x	Yanis	supergroup	688.89 KB	1	128 MB	FlumeData.1641038796915	
-rwxr-xr-x	Yanis	supergroup	660.72 KB	1	128 MB	FlumeData.1641038827950	
-rwxr-xr-x	Yanis	supergroup	650.65 KB	1	128 MB	FlumeData.1641038858925	
-rwxr-xr-x	Yanis	supergroup	676.85 KB	1	128 MB	FlumeData.1641038889933	
-rwxr-xr-x	Yanis	supergroup	671.69 KB	1	128 MB	FlumeData.1641038920894	

Maintenant que nos données ont été générées vers HDFS, nous pouvons maintenant les traiter avec HIVE.

Requêtes Hive

Le but de notre projet est de déterminer les raisons pour lesquelles les gens sont pour le vaccin et d'autres contre.

Twitter fournit des données au format JSON. Or Hive ne sait pas les traiter nativement. Il faut donc récupérer une librairie java qui va permettre à l'outil Hive de travailler avec ce format spécifique.

```
hive> ADD JAR json-serde-1.3-jar-with-dependencies.jar;
```

On va ensuite créer une table Hive. Une table externe car nous n'avons pas besoin de déplacer physiquement les données. Celles-ci resteront toujours sur HDFS et ne seront pas dupliquées.

```
hive> CREATE EXTERNAL TABLE IF NOT EXISTS ProVax (id string , user_friends_count int , user_location string , user_description string, user_statuses_count int, user_followers_count int , user_name string, user_screen_name string, created_at string , text string, retweet_count int , retweeted boolean , in_reply_to_user_id int, source string , in_reply_to_status_id int, media_url_https string, expanded_url string)
. . > ROW FORMAT DELIMITED FIELDS TERMINATED BY ',';
```

Puis, on importe les données depuis HDFS vers HIVE :

```
hive> LOAD DATA INPATH '/user/Hadoop/twitter_data/ProVax/*'
. . > OVERWRITE INTO TABLE ProVax;
Loading data to table default.provax
OK
```

Nous avons sélectionné 50 tweets PROVAX et 50 tweets ANTIVAX., que nous avons mis dans des tables intermédiaires pour nous faciliter le traitement

```
hive> INSERT INTO ProVax2 SELECT * FROM ProVax LIMIT 50 ;
```

On refait ensuite la même chose pour la partie ANTIVAX .

Les requêtes ainsi que leur résultats auxquels nous sommes parvenu sont comme suit:

Requêtes ANTIVAX

- SELECT user_screen_name,text FROM antivax2 where text like '%recul%';

```
hive> SELECT user_screen_name,text FROM antivax2 where text like '%recul%';
menalahy Je ne suis pas un #AntiVax mais je m'oppose au #Vaccin contre le #Covid_19 car il a été fabriqué
trop rapidement et qu'on n'a pas de recul. Les Défenseurs des #VaccinsAntiCovid_19 ont beau nous rassurer q
ue la Technologie et la Science ont évolué
STEPHILCAT Je ne suis pas #antivax je suis contre l'obligation d'une injection toujours en phase de tests
ou de nbx effets secondaires multiplient chaque jour ou personne n'ait assez de recul pour affirmer que c'est la
solution pour enrayer ce virus à alors non je ne me plierai pas
Dedel9177 Pour ma part j'ai bossé sans protection en 03/2020 je n'ai pas attrapé le Microbe ouf Je fais tou
s les gestes je ne fais pas la bise j'enlace pas mes parents ni amis je nettoie les emballages de courses e
tc Et je suis contre le vaccin car pas de recul sur les années à suivre
CapelucheR Ben moi dès que je touche mon smic je met 200 euros direct de côté c'est mon budget cigarette p
our le mois Je dis oui à la cigarette par contre je suis contre le vaccin car même si c'est gratuit on n'a
pas assez de recul et on sait pas ce qu'ils ont mis dedans
Brindacier Merci mais je suis contre ce vaccin car pas de recul et j'aurais toujours un doute pour ma fil
le Comment peut-on injecter des produits expérimentaux sans études préliminaires j'ai peur
5 rows selected (0,232 seconds)
```

- SELECT user_screen_name,text FROM antivax2 where text like '%mouton%';

```
hive> SELECT user_screen_name,text FROM antivax2 where text like '%mouton%';
Praadasoul J'ai dit que je me vaccinerai pas et je le ferai pas On est pas des moutons
johnwood68 Moi je ne vais pas me faire vacciner car mon corps m'appartient n'appartient pas au gouvernemen
t sur ce bon courage au petit mouton
2 rows selected (0,231 seconds)
```

- SELECT user_screen_name,text FROM antivax2 where text like '%libre%';

```
hive> SELECT user_screen_name,text FROM antivax2 where text like '%libre%';
anonymoumagg Mais ça je suis d'accord que chacun est libre de ses opinions de prendre ou non le vaccin Je
suis contre le passe et ceux qui le cautionnent car c'est nous enlever notre droit de choisir librement ce
qui est bon pour notre santé et notre droit à l'égalité
A_C_officiel Je ne me ferai pas vacciner car je n'ai pas envie d'être comme un gentil toutou à la politici
ue gouvernementale Je ne suis ni bête ni discipliné Je suis libre de choisir ce qui est bien pour mon corps
Ni #PassSanitaire ni quelconque intimidation ne changeront mon avis #Covid
NathalieBoisse7 Je ne me ferai pas vacciner car j'ai un très bon système immunitaire et je suis libre de
choisir ce qui est bon pour moi et mes enfants
AnnickGeffray Ni méfiante ni contestataire je ne me ferai pas vacciner car je suis encore libre libre de
dire NON à leur injection politique plus que sanitaire
Michael80361703 Toujours à 0 dose bientôt 65 ans je trouve que le pass vaccinal est encore plus scandaleu
x que le pass sanitaire Je ne me ferai pas vacciner Je suis libre car je choisis de ne pas me laisser impos
er une injection d'un produit inconnu car ce n'est pas un vaccin
ReineMargot2 Je ne me ferai jamais vacciner car je suis libre je n'ai pas été élevé dans du coton j'ai
grandi dans la rébellion je m'en suis sortie plus forte et maintenant je devrais OBEIR Jamais Je refuse d
e vivre sous contrôle je refuse que des gens décident pour moi
6 rows selected (0,26 seconds)
```

- SELECT user_screen_name,text FROM antivax2 where text like '%confiance%';

```
hive> SELECT user_screen_name,text FROM antivax2 where text like '%confiance%';
xRisYn La situation complète a été conçue pour ça Virus ou pas Les gens pensent ce qu'ils veulent mais l'ét
at a fait le choix de diviser Je suis contre le vaccin car je n'ai pas confiance mais j'aurai plutôt une o
bligation Ou le choix sans restriction parmi d'autres moyens
CavigliaLauren1 Bonjour Je suis sûr par vos propos Je ne me ferai pas vacciner car je n'ai aucune confi
ance en vous en la parole du gouvernement Ce n'est bien sûr pas la raison principale mais c'est celle qui e
nfonce le clou
2 rows selected (0,23 seconds)
```

- SELECT user_screen_name,text FROM antivax2 where text like '%peur%';

```
hive> SELECT user_screen_name,text FROM antivax2 where text like '%peur%';
ddg7893 Je n'ai pas peur de le dire je suis #antivax
le_yann_pi Je suis contre le vaccin car j'ai peur des aiguilles
Aniel9711985 @PascalPraud Je ne me ferai pas vacciner car j'ai peur de mourir d'une thrombose Le vaccin me
fait plus peur que le virus De plus je ne crois plus les soi-disant experts car on s'est aperçu qu'ils éta
ient eux même perdus
Brindacier Merci mais je suis contre ce vaccin car pas de recul et j'aurais toujours un doute pour ma fil
le Comment peut on injecter des produits expérimentaux sans études préliminaires j'ai peur
4 rows selected (0,277 seconds)
```

- SELECT user_screen_name,text FROM antivax2 where text like '%cobaye%';

```
hive> SELECT user_screen_name,text FROM antivax2 where text like '%cobaye%';
fmbangoura7 A rien du tout Moi je ne me vaccine pas Raising hands Je suis pas un cobaye yoo
1 row selected (0,218 seconds)
```

- SELECT user_screen_name,text FROM antivax2 where text like '%risque%';

```
hive> SELECT user_screen_name,text FROM antivax2 where text like '%risque%';
jacquesderosier Cela dit je suis 0 risque mais je ne me ferai pas vacciner tant qu'il y aura cette ambian
e d'ultrapre Je précise que je ne suis pas antivax Mon choix est strictement personnel
DelormeEric2 J'ai fait le calcul de probabilité j'ai moins de risque de mourir du covid que de développer
une maladie rare après vaccination
2 rows selected (0,237 seconds)
```

- SELECT user_screen_name,text FROM antivax2 where text like '%crois%';

```
hive> SELECT user_screen_name,text FROM antivax2 where text like '%crois%';
Aniel9711985 @PascalPraud Je ne me ferai pas vacciner car j'ai peur de mourir d'une thrombose Le vaccin me
fait plus peur que le virus De plus je ne crois plus les soi-disant experts car on s'est aperçu qu'ils éta
ient eux même perdus
1 row selected (0,234 seconds)
```

- SELECT user_screen_name,text,created_at FROM antivax2 where created_at like '%Dec%' AND created_at like '%2021%';

```
hive> SELECT user_screen_name,text,created_at FROM antivax2 where created_at like '%Dec%' AND created_at li
ke '%2021%';
Herve de Coren pas de gosse juste des chats mais celui qui voudrai vacciner mon chat je le castrer avec le
dents Next time pour les invalides immunodéprimés comme suppression de la pension donc de toute ressource
et JE NE ME FERAI PAS VACCINER car mon consentement n'a pas été éclairé Mon Dec 20 03:08:00 +0000 2021
77Calixte tant enceinte voilà pourquoi je ne me ferai pas vacciner Mon Dec 27 19:01:00 +0000 2021
jacquesderosier Cela dit je suis 0 risque mais je ne me ferai pas vacciner tant qu'il y aura cette ambian
e d'ultrapre Je précise que je ne suis pas antivax Mon choix est strictement personnel Sun Dec 26 11:33:00
+0000 2021
AnnickGeffray Ni méfiante ni contestataire je ne me ferai pas vacciner car je suis encore libre libre de
dire NON à leur injection politique plus que sanitaire Wed Dec 29 18:23:00 +0000 2021
Sarah41659531 Personnellement je n'accepterai aucune extorsion de consentement Je ne me ferai pas vaccine
avec un produit douteux et qui entraîne une addiction même pour aller à un RV à l'hôpital C'est trps clai
dans ma tête Pas d'injection FORCÉE Surtout si ça ne fonctionne pas Tue Dec 28 09:55:00 +0000 2021
Michael80361703 Toujours 0 dose bientôt 65 ans je trouve que le pass vaccinal est encore plus scandaleu
x que le pass sanitaire Je ne me ferai pas vacciné Je suis libre car je choisis de ne pas me laisser imposer
une injection d'un produit inconnu car ce n'est pas un vaccin Mon Dec 27 20:38:00 +0000 2021
ReineMargot2 Je ne me ferai jamais vacciner car je suis libre je n'ai pas été élevée dans du coton j'ai
grandi dans la rébellion je m'en suis sortie plus forte et maintenant je devrais OBEIR Jamais Je refuse
de vivre sous contrôle je refuse que des gens décident pour moi Sat Dec 18 09:43:00 +0000 2021
grandosek 0.05 % = le taux de létalité Moyenne d'âge des décès 82 ans Tests biaisés car amplifiés jusqu'à
45 fois Il y a un traitement qui marche si pris à temps Les masques se portent au bloc Vaccin non testés s
ur des animaux Pour toutes ces raisons je ne me ferai pas vacciner Fri Dec 24 13:09:00 +0000 2021
DelormeEric2 J'ai fait le calcul de probabilité j'ai moins de risque de mourir du covid que de développer
une maladie rare après vaccination Fri Dec 24 12:35:00 +0000 2021
laurentAPERO Non je ne ferai pas vacciné car c'est pas un vaccin la définition du mot vaccin et ça ne cor
respond rien au produit injecté Substance pathogène qui inoculée à un individu lui confère l'immunité co
tre une maladie Donc on vit normalement sans ce produit en nous Thu Dec 30 14:28:00 +0000 2021
23 rows selected (0,264 seconds)
```

- SELECT user_screen_name,text,created_at FROM antivax2 where created_at like '%Jul%' AND created_at like '%2021%';

```
hive> SELECT user_screen_name,text,created_at FROM antivax2 where created_at like '%Jul%' AND created_at li
ke '%2021%';
Dedel9177 Pour ma part j'ai bossé sans protection en 03/2020 je n'ai pas attrapé le Microbe ouf Je fais tj
s les gestes je ne fais pas la bise j'enlace pas mes parents ni amis je nettoie les emballages de courses e
tcà Et je suis contre le vaccin car pas de recul sur les années à suivre Wed Jul 21 16:29:00 +0000 2021
le_yann_pi Je suis contre le vaccin car j'ai peur des aiguilles Thu Jul 15 09:08:00 +0000 2021
johnwooo000 Moi je ne vais pas me faire vacciner car mon corps m'appartient n'appartient pas au gouvernemen
t sur ce bon courage au petit mouton Tue Jul 13 17:37:00 +0000 2021
freezaw_ Tu est le meilleur perso je ne vais pas me faire vacciner car je n'en ai pas le besoin Mais je co
nstate que notre gouvernement osé toucher au loi des droits de l'homme pour notre détruire notre liberté En
fin je vais avoir 80 000 commentaires toxique à m'insulter en dessus Tue Jul 13 08:45:00 +0000 2021
Parisien92aze Ah mais entièrément d'accord Je ne remets absolument pas en cause l'efficacité du vaccin à c
ourt terme les chiffres sont clairs Je ne me ferai pas vacciner car il n'y a en revanche aucune étude sur l
es effets à long terme et j'ai pas prouvé de me faire piquer tous les 6 mois Fri Jul 30 16:28:00 +0000 202
1
Caviglialauren1 Bonjour Je suis étonné par vos propos Je ne me ferai pas vacciner car je n'ai aucune confi
ance en vous en la parole du gouvernement Ce n'est bien sûr pas la raison principale mais c'est celle qui e
nfonce le clou Sat Jul 17 15:48:00 +0000 2021
NathalieBoisse7 Je ne me ferai pas vacciner car j'ai un très bon système immunitaire et je suis libre de
choisir ce qui est bon pour moi et mes enfants Sun Jul 11 01:42:00 +0000 2021
7 rows selected (0,236 seconds)
```

Requetes PROVAX

- SELECT user_screen_name,text FROM provax2 where text like '%prot%' or text like '%proteger' ;

```
hive> SELECT user_screen_name,text FROM provax2 where text like '%prot%' or text like '%proteger' ;
ThomasScordia C'est pas pour moi que je me suis fait vacciné car pas spécialement à risque mais pour proté
ger les autres qui peuvent l'être
AlainMartin458 C vous qui écrivez des conneries Le vaccin protège et il faut se vacciner car il n'y a pas
mieux
Panam_inside C'est un double voir triple discours on demande aux vaccinés de se faire tester pour pas prop
ager le virus mais on te dit aussi qu'il faut se faire vacciner car sinon on en viendra pas à bout mais le
vaccin empêche la propagation ou pas Et c'est pour protéger qui en fait
Maximelg21 InfirmierVaccine depuis 6 mois et tout va bien Je suis pour la vaccination car c'est le moyen l
e plus efficace pour me protéger ainsi que mes proches et mes patients Par contre si tu es encore indécis j
e ne te jugerai pas Viens on discute plutôt Visage faisant un clin d'œil
SerotineN Je suis pour Simplement parce que je n'étais pas contre les autres vaccins Ça protège un vaccin
anthoneymoon Après moi je suis pour le vaccin car ça protège des formes graves on peut pas le nier Mais à
partir du moment où tu fais des tests régulièrement je ne vois pas où est le soucis
6 rows selected (0,239 seconds)
```

- SELECT user_screen_name,text FROM provax2 where text like '%mieux%';

```
hive> SELECT user_screen_name,text FROM provax2 where text like '%mieux%';
Sauvage35249696 Je suis provax car les chiffres montrent clairement l'action positive du vaccin et il fa
ut mieux prévenir que guérir La vaccination est très peu risquée mais elle ne semble pas résoudre l'ensembl
e du problème
AlainMartin458 C vous qui écrivez des conneries Le vaccin protège et il faut se vacciner car il n'y a pas
mieux
MurielleLombard je suis pro vax car on a rien de mieux pour le moment
3 rows selected (0,248 seconds)
```

- SELECT user_screen_name,text FROM provax2 where text like '%voyage%' or text like '%vacance%';

```
hive> SELECT user_screen_name,text FROM provax2 where text like '%voyage%' or text like '%vacance%';
LeaAssayag Je me suis vacciné seulement pour partir en vacance c'est tout et parce que les tests antigéniqu
e font trop mal
pascalinho11 Je me suis vacciné car entre le boulot le foot (payer pour un viol nasal toutes les semaines
pas d'intimité) et les voyages en Italie pas le choix Franchement à part une douleur au bras pendant 2 jours
rien à signaler
SB3141 Moi je me suis vacciné car tuer des gens ça me plaît moyen Les vacances c'est du bonus
3 rows selected (0,229 seconds)
```

- SELECT user_screen_name,text FROM provax2 where text like '%test%';

```
hive> SELECT user_screen_name,text FROM provax2 where text like '%test%';
jdcsef Je me suis fait vaccin  car je crois dans la science et j'ai m me fait partie d'un test clinique pr
o deo  videment d'un vaccin non homologu  pour prendre ma part et augmenter le panel de solution contre cet
te  pid mie
leaAssayag Je me suis vaccin  seulement pour partir en vacance c est tout et parce que les test antigeniq
ue font trop mal
agmlscallaert Je suis vaccin  car je crois en la vaccination Certains sont vaccin s pour pouvoir  tre libr
e d aller ou ils veulent D autre l a refuse encore Et l  on apprend qu un test neg sera obligatoire en plus
du pass vaccinale dans certain lieu Sa sera sans moi cette fois
florent_57 Et pour avoir le pass il faut se faire VACCINER car aucun non vaccin  ne pourra payer un test t
out les 3 jours pour faire ses courses
Panam_inside C'est un double voir triple discours on demande aux vaccin s de se faire tester pour pas prop
ager le virus mais on te dis aussi qu'il faut se faire vacciner car sinon on en viendra pas a bout mais le
vaccin emp che la propagation ou pas Et c'est pour prot ger qui en fait
anthoneymoon Apr s moi je suis pour le vaccin car  a prot ge des formes graves on peut pas le nier Mais  
partir du moment o  tu fais des test r gul rement je ne vois pas o  est le soucis
6 rows selected (0,252 seconds)
```

- SELECT user_screen_name,text FROM provax2 where text like '%croi%';

```
hive> SELECT user_screen_name,text FROM provax2 where text like '%croi%';
jdcsef Je me suis fait vaccin  car je crois dans la science et j'ai m me fait partie d'un test clinique pr
o deo  videment d'un vaccin non homologu  pour prendre ma part et augmenter le panel de solution contre cet
te  pid mie
agmlscallaert Je suis vaccin  car je crois en la vaccination Certains sont vaccin s pour pouvoir  tre libr
e d aller ou ils veulent D autre l a refuse encore Et l  on apprend qu un test neg sera obligatoire en plus
du pass vaccinale dans certain lieu Sa sera sans moi cette fois
shinyLucille Un ancien m a dit qu il a trop vu dans sa jeunesse des personnes devenir handicap es moteur e
t mentales   cause de la polio  a a fait des ravages donc depuis il est pro vaccin Je suis pro vaccin car j
e crois en la m decine pour sortir de ce cauchemar
3 rows selected (0,24 seconds)
```

- SELECT user_screen_name,text FROM provax2 where text like '%conf%';

```
hive> SELECT user_screen_name,text FROM provax2 where text like '%conf%';
Nihm1313 C est scandaleux Je suis pour le vaccin car j ai confiance dans l avanc e de la m decine et je su
is contre le pass sanitaire car j ai confiance dans les fran ais Mais non il a fallu qu il impose interdise
Voil  comment Emmanuel Macron rate la concorde avec ses citoyens
raguhnna Je suis pro vaccin car je suis scientifique Et qu aucun scientifique de bon sens n est contre La s
ituation avec mon fils a bien confirm  l int r t de la vaccination
2 rows selected (0,238 seconds)
```

- SELECT user_screen_name,text FROM provax2 where text like '%libre%' OR text like '%avant%';

```
hive> SELECT user_screen_name,text FROM provax2 where text like '%libre%' OR text like '%avant%';
thibzpower Le Vaccin est le seul moyen pour lutter contre le virus et retrouver sereinement la vie d'avant
Donc les antivax si vous pensez vous rebeller contre le gouvernement en n'allant pas vous faire vacciner n
on vous  tes juste  go stes et vos arguments sont pitoyables
matteooklm Je veux que les choses reviennent comme avant donc forc ment  a passe par le vaccin + se faire
vacciner c'est pas la mort mdr 
agmlscallaert Je suis vaccin  car je crois en la vaccination Certains sont vaccin s pour pouvoir  tre libr
e d aller ou ils veulent D autre l a refuse encore Et l  on apprend qu un test neg sera obligatoire en plus
du pass vaccinale dans certain lieu Sa sera sans moi cette fois
victoriahpln 2022 Je vais reprendre mon r gime et mon sport m me me vaccin  pour aller   la salle ma vie d
 avant me manque
Franklin_6_N Je suis pour le vaccin car j'aimerais avoir une vie comme avant La maladie en elle m me ne me
fait pas peur Je veux juste  tre de nouveau libre de mes mouvements ici en Europe
5 rows selected (0,226 seconds)
```

- SELECT user_screen_name,text FROM provax where text like '%consci%' OR text like '%convaincu%' ;

```
hive> SELECT user_screen_name,text FROM provax where text like '%consci%' OR text like '%convaincu%' ;
Mister576 On doit agir en pleine conscience et faire ce que notre raison nous ordonne tout en respectant l
es avis des autres Je me suis fait vaccin  car je suis convaincu de son efficacit  et rien ne m'y a oblig 
1 row selected (0,223 seconds)
```

- SELECT user_screen_name,text FROM provax2 where text like '%efficace%';

```
hive> SELECT user_screen_name,text FROM provax2 where text like '%efficace%';
Noisethino Pour le vaccin est le seul moyen de sortir de la crise il est efficace s'r et gratuit c'est une
chance de pouvoir l'avoir aussi rapidement
MaximeIg2l InfirmierVaccine depuis 6 mois et tout va bien Je suis pour la vaccination car c'est le moyen l
e plus efficace pour me protéger ainsi que mes proches et mes patients Par contre si tu es encore indécis j
e ne te jugerai pas Viens on discute plut'ôt Visage faisant un clin d'œil
2 rows selected (0,222 seconds)
```

- SELECT user_screen_name,text,created_at FROM provax2 where created_at like '%Dec%' AND created_at like '%2021%' or created_at like '%Oct%' or created_at like '%Nov%';

```
hive> SELECT user_screen_name,text,created_at FROM provax2 where created_at like '%Dec%' AND created_at lik
e '%2021%' or created_at like '%Oct%' or created_at like '%Nov%';
S_T_R_U Oui je suis prêt à me faire vacciner autant de fois qu'il le faudra en effet Et si la bêtise tue T
ous les jours des antivaccins meurent Sun Oct 05 10:07:08 +0000 2021
raguhna Je suis pro vaccin car je suis scientifique Et qu'aucun scientifique de bon sens n'est contre La s
ituation avec mon fils a bien confirmé l'intérêt de la vaccination Wed Nov 17 18:08:00 +0000 2021
victoriahpln 2022 Je vais reprendre mon régime et mon sport même me vaccin' pour aller à la salle ma vie d
'avant me manque Fri Dec 31 21:35:00 +0000 2021
cosmic_blue Oui pour réduire drastiquement mes chances de tomber malades et que le virus ne mute + à cause
de moi Ceci dit on est tous vaccinés depuis des années à d'autres maladies que le covid Sat Dec 18 22:31
:00 +0000 2021
florent_57 Et pour avoir le pass il faut se faire VACCINER car aucun non vacciné ne pourra payer un test t
out les 3 jours pour faire ses courses Sun Oct 05 10:07:08 +0000 2021
Panam_inside C'est un double voir triple discours on demande aux vaccinés de se faire tester pour pas prop
ager le virus mais on te dis aussi qu'il faut se faire vacciner car sinon on en viendra pas à bout mais le
vaccin empêche la propagation ou pas Et c'est pour protéger qui en fait Tue Dec 24 10:04:08 +0000 2021
Lorenzato47 C'est exactement pour cela qu'il faut se faire vacciner Car si tu as un super syst'ème immunita
ire et que tu lui apprends en plus à détruire le tu deviendras un trou noir à virus. Et tu sauveras malgr'
toi des vies Fri Dec 24 14:56:00 +0000 2021
S_K_D Je suis pour le vaccin car je suis contre l'idée qu'on restreigne ma liberté sur le long terme Or c'
est ce qui va se passer si on n'est pas vacciné Donc je ne comprends vraiment pas les antivax Leur logique
me semble tr'ès paradoxale Mon Nov 22 10:07:08 +0000 2021
francelibre1958 Je suis pour le vaccin car dans la plupart des cas ça marche j'ai eu mes 2 doses et 0 effe
ts secondaires Apr's je ne nie pas qu'il puisse y avoir des effets secondaires mais ils repr'sentent 0 Th
u Dec 23 15:31:00 +0000 2021
FabFabuleuse Normalement je suis pour le vaccin car de toute façon on n'a pas de médicament contre le cov
id perso je suis du genre à craindre le pire que ça met en quelque chose de pire mais avec tout ça on se po
se la question Visage avec les yeux levés au ciel car même le président personne ne l'a vu se faire vaccin'
ersonne haussant les épaules Sat Dec 25 10:55:00 +0000 2021
16 rows selected (0,256 seconds)
```

- SELECT user_screen_name,text,created_at FROM provax2 where created_at like '%Aug%' AND created_at like '%2021%' or created_at like '%Jul%' or created_at like '%Jun%';

```
hive> SELECT user_screen_name,text,created_at FROM provax2 where created_at like '%Aug%' AND created_at lik
e '%2021%' or created_at like '%Jul%' or created_at like '%Jun%';
```

```
du pass vaccinale dans certain lieu ça sera sans moi cette fois Thu Aug 19 15:07:08 +0000 2021
idriiss_neumann 100% en phase Perso je suis favorable au vaccin car de ce que j'en comprend le bénéfice ri
que est faible Par contre je le dit tout de suite je compte pas le faire tout les 3/4 mois même 1 fois par
ans c'est beaucoup Mon Jul 5 12:47:00 +0000 2021
SB3141 Moi je me suis vaccin' car tuer des gens ça me pla't moyen Les vacances c'est du bonus Thu Jul 23
14:30:00 +0000 2021
Charles_Chauve Je suis pour la vaccination car c'est le seul moyen connu d'enrayer une pandémie Et nous
avons de nombreux exemples qui le prouvent Et je suis vaccin' contre le Covid 19 MAIS je suis également pou
l'obligation du casier judiciaire vierge pour tous les Hommes politiques Mon Jul 05 07:48:00 +0000 2021
MaximeIg2l InfirmierVaccine depuis 6 mois et tout va bien Je suis pour la vaccination car c'est le moyen l
e plus efficace pour me protéger ainsi que mes proches et mes patients Par contre si tu es encore indécis j
e ne te jugerai pas Viens on discute plut'ôt Visage faisant un clin d'œil Thu Jul 08 07:49:00 +0000 2021
Rue42919748 J'ai pas dit ça je suis pour la vaccination car celle-ci nous permet de vivre plus longtemps
e met en lumi'ère le fait que peut importe ce que l'on fait les virus seront tjr de plus en plus fort Fri
Jul 30 18:45:08 +0000 2021
LZagorac Oui tu as raison le débat embrasse en réalité tout notre mode de vie et de développement qui d'
p'le l'écossystème global Je suis pour la vaccination car je préfère prévenir que guérir mais prévenir r'ell
ement c'est créer les conditions pour ne plus développer de zoonoses Sat Aug 07 09:17:00 +0000 2021
Reisendenn Perso je suis pro vax Car quand on se renseigne suffisamment auprès de médecins et de sp'cialis
tes du domaine quand on comprend comment ça fonctionne il n'y a aucune raison de refuser ce vaccin Surtout
les ARNm Mais ceux qui doutent sont juste mal informés ça pourrait Tue Jul 13 12:31:00 +0000 2021
MurielleLombard je suis pro vax car on a rien de mieux pour le moment Thu Jul 29 11:07:00 +0000 2021
20 rows selected (0,262 seconds)
```

Analyse des résultats

A travers ce projet, nous avons pu voir les raisons pour lesquelles les Français étaient pour ou contre le vaccin du covid-19.

Les raisons pour lesquelles les Français sont pour le vaccin sont les suivantes :

- Se protéger contre les différents variants surtout ceux qui sont sujets à risque.
- Minimiser la propagation du virus.
- Voyager/ partir en vacances.
- Éviter de se faire tester à chaque fois grâce au pass sanitaire.
- Faire confiance aux chercheurs.
- Croire à la science et à l'efficacité du vaccin.
- Être libre et retrouver la vie d'avant.
- Être conscient.

Nous avons également pu conclure que les français se manifestaient sur twitter pour défendre leur position de pro-vaccin beaucoup plus en mois de Décembre, ce qui s'explique avec l'arrivée de la deuxième vague après les fêtes de fin d'année, mais aussi en été avec l'arrivée des vacances.

Nous allons maintenant voir les raisons pour lesquelles les Français sont contre le vaccin :

- Ne pas avoir assez de recul sur le vaccin.
- Se sentir comme des moutons ou des cobayes.
- Avoir peur des effets secondaires que pourraient engendrer le vaccin.
- Avoir peur de mourir.
- Ne pas se sentir libre.
- Ne pas avoir confiance en la substance qui sera injectée dans les corps.

Conclusion

Dans ce projet, Hadoop nous a permis d'analyser des données non structurées et grâce à cela nous avons pu analyser les tweets de certaines personnes dans le thème de la vaccination contre la covid-19.

Nous avons donc pu voir les raisons pour lesquelles une partie de la population s'oppose à la vaccination et une autre partie est pour.

Ainsi, cela nous permet de mieux comprendre les différents points de vue de chacun.