

Donalds Trump all tweets data

Charley GEOFFROY

2023-04-26

Donald Trump Tweets Dataset

Dans le cadre de cette analyse, nous avons choisi d'étudier les tweets de Donald Trump, une personnalité politique qui a suscité de nombreuses polémiques à travers ses messages sur les réseaux sociaux. Les tweets de Donald Trump ont souvent fait parler de lui et ont été l'un des principaux moyens par lesquels il a communiqué avec le public et exprimé ses opinions.

Nous avons décidé de nous pencher sur ce sujet car les tweets de Donald Trump ont eu un impact significatif sur le paysage politique, médiatique, mais aussi **démocratique**. Il convient de noter que Donald Trump a été confronté à de nombreuses controverses liées à ses tweets et à leur contenu parfois provocateur et virulent.

Dans notre quête de compréhension, il est également crucial d'explorer qui se cache réellement derrière les tweets de Donald Trump. Alors que le compte Twitter officiel de Donald Trump était géré par son équipe de communication, il était bien connu que certains tweets étaient rédigés directement par Donald Trump lui-même. Cette distinction entre les tweets rédigés par le président lui-même et ceux rédigés par son équipe est importante pour analyser les différents aspects et styles de communication présents dans ses tweets.

En outre, nous chercherons également à comprendre de manière factuelle pourquoi et comment certains tweets de Donald Trump ont été supprimés.

Présentation du jeu de données

Attention Les données utilisées proviennent du site **Kaggle** et ont été créées par **@Shantanu Roy**. Il est important de noter que bien que l'auteur ait pris des mesures pour garantir la conformité et la véracité des données, leur intégralité ne peut être entièrement garantie. Veuillez prendre cela en considération lors de l'analyse et de l'interprétation des résultats obtenus à partir de ces données.

Veuillez trouver le data set ici

Avant de commencer notre analyse statistique, permettez-moi de vous présenter quelques chiffres clés qui contribueront à une meilleure compréhension des résultats et qui appuient factuellement la montée en notoriété de Donald Trump sur le réseau social Twitter :

Donal Trump, c'est un total de :

- Tweets : **56,571**
- Tweets supprimés : **1,092**
- Signalés : **304**
- Retweeté par sa communauté: **487,584,740**

Présentation de l'évolution des tweets de Donald Trump par année

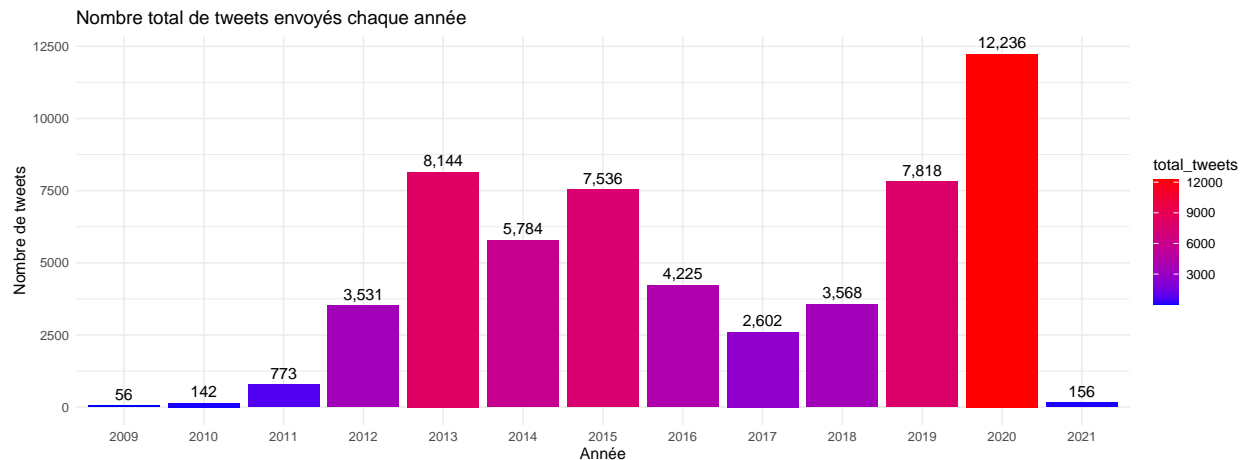
Question: Combien de tweets a envoyé Donald Trump pour chaque année ?

Individu: un Tweet de Donald Trump

Population: Tous les tweets de Donald Trump par année

Échantillon: Kaggle Dataset Donald Trump tweets de 2009-2021

Variable: Nombre total de tweets, quantitative discrète



L'histogramme présentant le nombre total de tweets envoyés par Donald Trump révèle deux moments clés dans son utilisation de Twitter.

Entre 2013 et 2015, il a été extrêmement actif sur la plateforme, envoyant jusqu'à **8,144 tweets en 2013**. Cependant, le nombre de tweets envoyés a drastiquement **diminué entre 2016 et 2018**, ce qui correspond à une partie de son mandat présidentiel, passant de **7,536 en 2015** à seulement **4,225 en 2016**.

Ensuite, à partir de 2019 jusqu'à la fin de son mandat présidentiel en 2020, Donald Trump est revenu en force sur Twitter, le nombre de tweets envoyés explosant à nouveau avec **7,818 tweets en 2019** et **12,236 tweets en 2020**.

Ces fluctuations dans l'utilisation de Twitter par Donald Trump peuvent être attribuées à des événements politiques, tels que des périodes électorales où il a envoyé un grand nombre de tweets pour faire parler de lui.

Corrélation entre le nombre total Tweets Flagged et Deleted

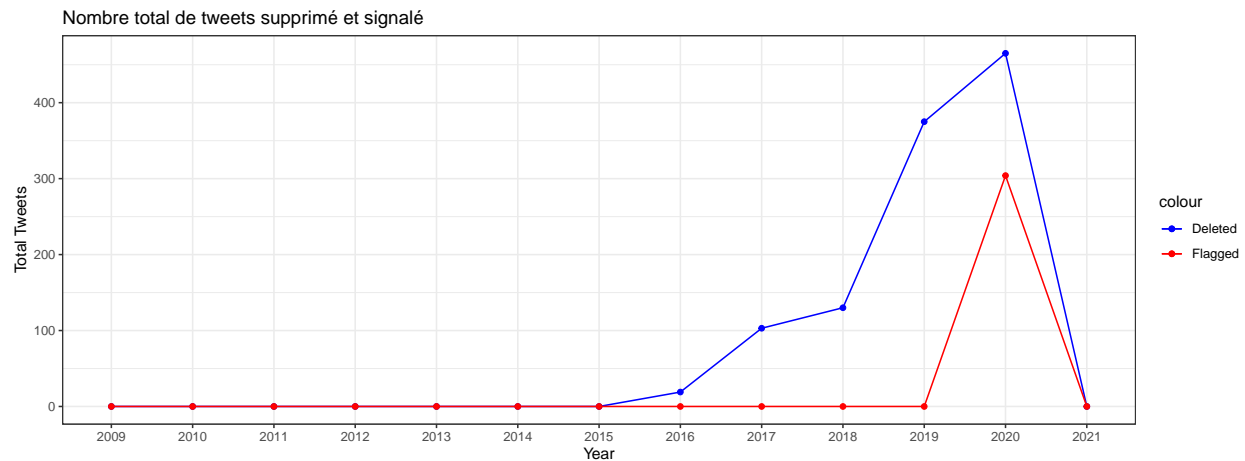
Question: Existe-il une corrélation entre le nombre total de tweets supprimés de Donald Trump et le nombre total de tweets signalés?

Individu: un Tweet de Donald Trump supprimé, un Tweet signalé

Population: L'ensemble des tweets supprimés et signalés de Donald Trump de 2009 à 2017 et de 2017 à 2021.

Échantillon: Kaggle Dataset Donald Trump tweets de 2009-2021

Variable: Nombre total de tweets supprimé et le Nombre total de tweets signalés, quantitative discrète



Il est intéressant de noter que **le nombre de tweets supprimés de Donald Trump est plus élevé que le nombre de tweets signalés**. Cette disparité peut sembler paradoxale, car généralement on s'attendrait à ce que les signalements soient plus fréquents que les suppressions. Cela soulève des questions quant à la transparence du processus de modération des contenus sur la plateforme et peut susciter des doutes sur l'existence d'une possible censure ou d'un traitement préférentiel.

Cependant, il est important de souligner que ces données brutes ne fournissent pas suffisamment d'informations contextuelles pour tirer des conclusions définitives. Il pourrait y avoir des explications légitimes à cette disparité, telles que des politiques spécifiques de modération mises en place par Twitter ou des différences dans le traitement des signalements.

Répartition des tweets de Donald Trump par device

Question: Combien de tweets poste Donald Trump par device pour chaque année ? Qu'elles sont les devices représentatif par période ?

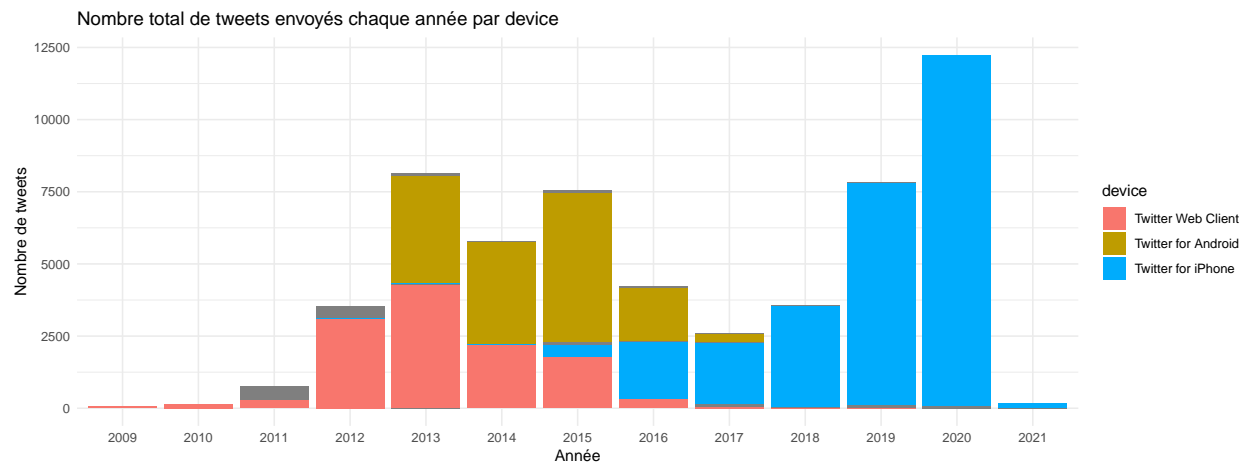
Individu: un Tweet de Donald Trump

Population: L'ensemble des tweets de Donald Trump.

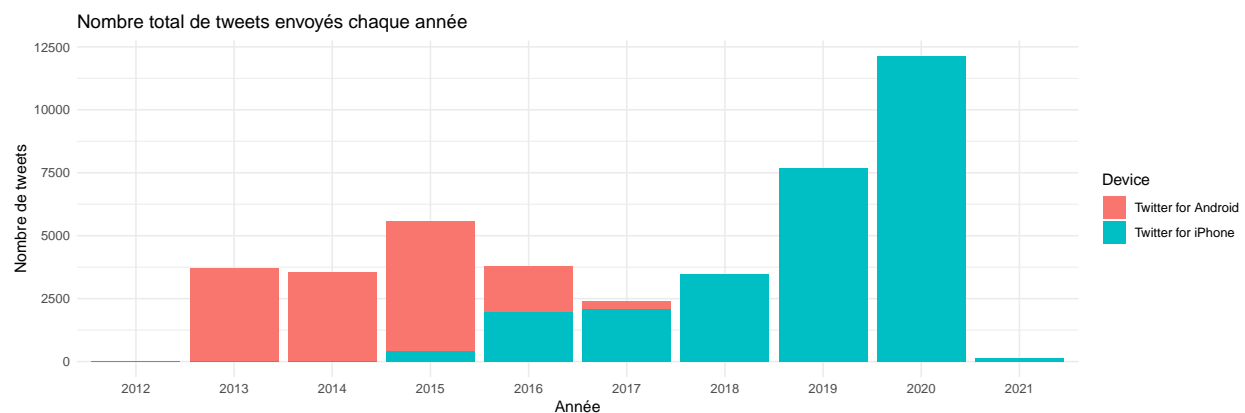
Échantillon: Kaggle Dataset Donald Trump tweets de 2009-2021

Variable && modalité: Nombre de tweets envoyé par device, Nombre de device majoritaire, quantitative discrète

Nombre de tweet envoyé par devices (> 500 fois)



Répartition en fonction des devices android et iphone



D'après les deux histogrammes, **plusieurs types d'appareils ont été utilisés** pour publier sur le compte Twitter de Donald Trump. Lorsque nous filtrons les données pour ne récupérer que les appareils ayant publié plus de 500 fois, nous constatons qu'au total **4 appareils sont présents : l'iPhone, Android, le client web Twitter et un appareil inconnu**.

En Janvier 2017 pour fêter l'arrivée de Donald Trump en tant que président, le magazine The Guardian a rapporté que celui-ci utilisait un Samsung Galaxy S3, téléphone considéré comme non sécurisé.

En examinant de plus près le deuxième diagramme qui ne montre que les appareils iPhone et Android, nous pouvons déduire deux choses de ces informations corrélées à celle du magazine **THE GUARDIAN** :

Nous constatons un début de publication à partir d'un iPhone en 2015, ce qui pourrait correspondre à **l'arrivée de l'équipe de communication** de Donald Trump pour les élections car à cette époque Donald Trump utilisait un Android.

Nous constatons aussi, qu'en 2017, malgré les risques de vulnérabilité, **un téléphone Android continuait de publier** sur le compte Twitter mais que la majorité des tweets sont des devices iPhones.

Ainsi, peut-on déduire que les tweets envoyés depuis un iPhone de 2015 à 2016 ont été envoyés par l'équipe de communication de Donald Trump ? De plus, est-ce que la mise en cause de l'utilisation d'un téléphone

non sécurisé dans l'article du Guardian pourrait expliquer l'arrêt soudain des publications sous Android en 2017 ?

Tendance du nombre de tweets envoyé & nombre de tweets supprimés par heure de Donald Trump

Questions: Quelle est la tendance du nombre de tweets envoyé par heure, Les tweets supprimés de Donald Trump suivent-ils une tendance similaire ? Que peut-on en déduire ?

Individu: un Tweet de Donald Trump, un tweet supprimé de Donald Trump

Population: L'ensemble des tweets de Donald Trump de 2009-2017.

Échantillon: Kaggle Dataset Donald Trump tweets de 2009-2021

Variable && modalité: Nombre de tweets envoyé par heure, Nombre tweets supprimé par heure, quantitative discrète

HISTOGRAMME

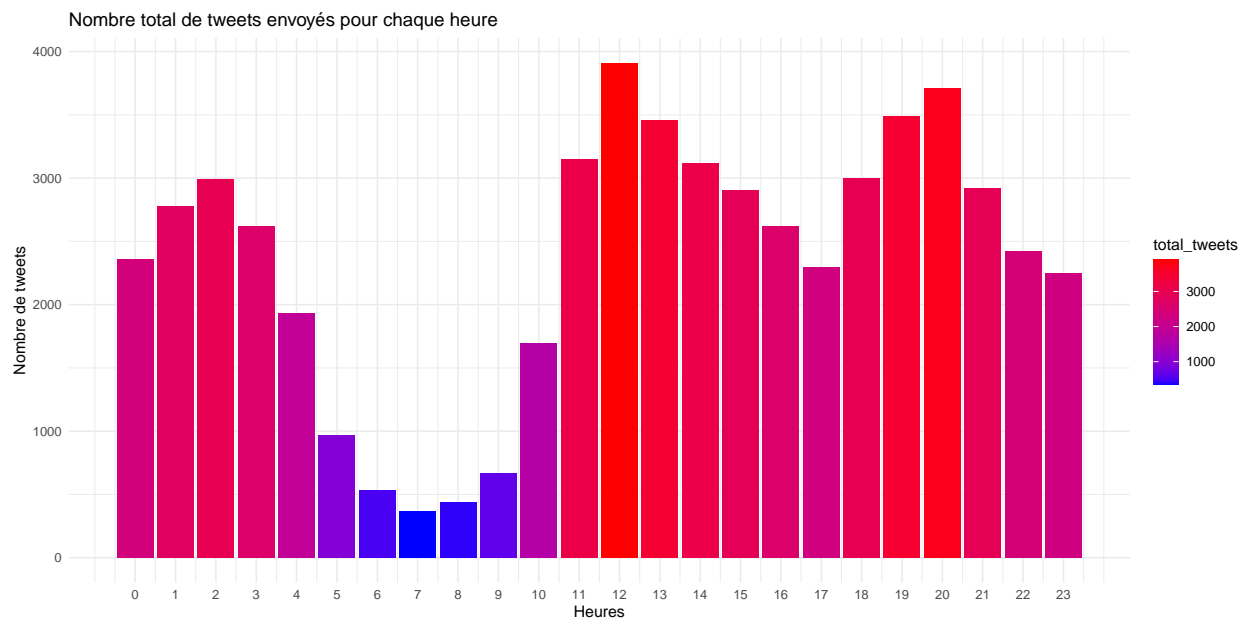
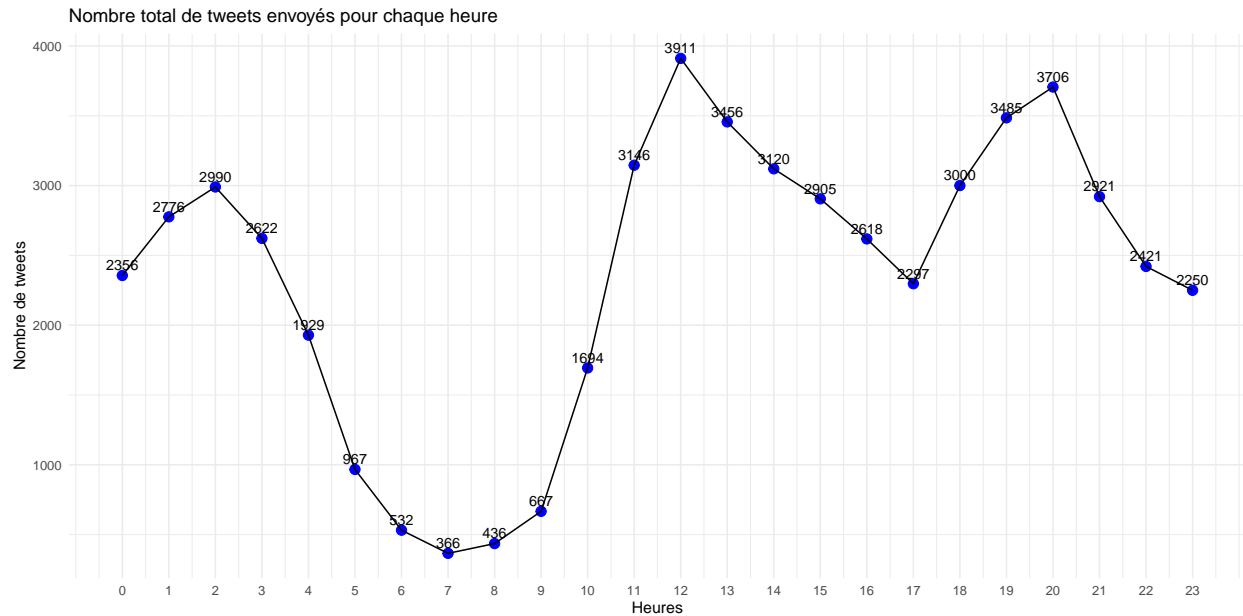


DIAGRAMME NUAGE DE POINTS

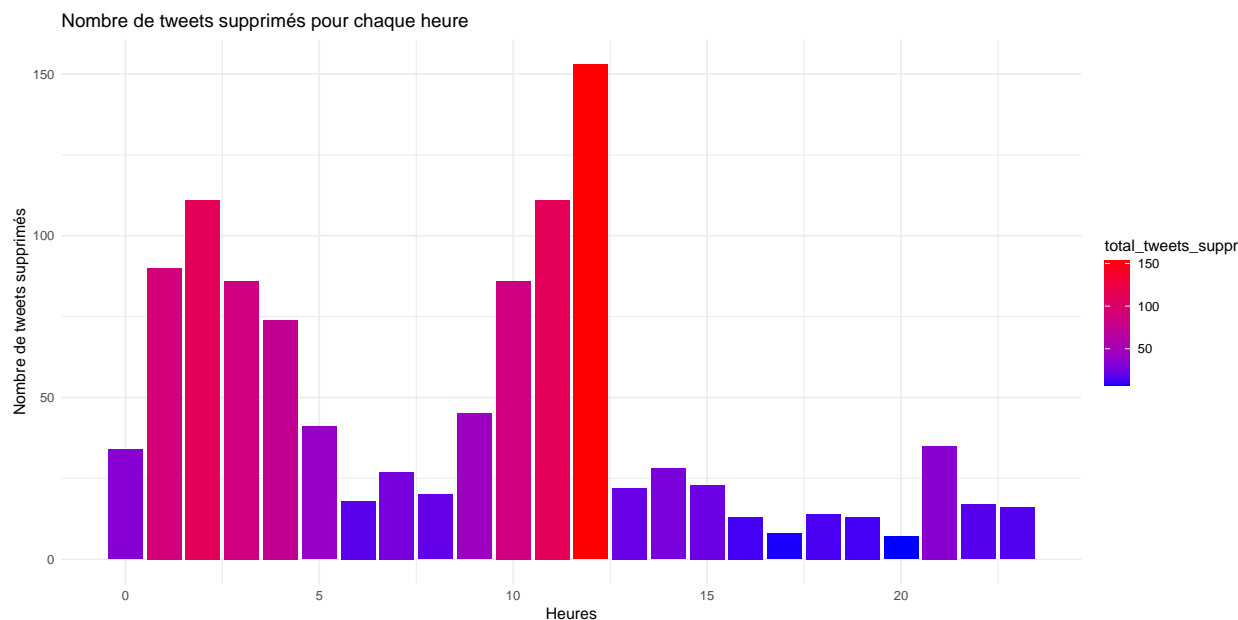


L'analyse de l'histogramme et du nuage de points nous permet de déduire plusieurs éléments intéressants. Tout d'abord, il apparaît que le compte de Donald Trump est le plus actif **autour de midi**, avec un pic atteignant un total de **3911 tweets envoyés**. Ensuite, le nombre de tweets **diminue progressivement jusqu'à environ 17 heures**, où il atteint **un total de 2297 tweets**. Par la suite, il augmente à nouveau progressivement pour atteindre **3706 tweets vers 20 heures**.

Il est également notable que Donald Trump publie très peu entre 5 heures et 9 heures du matin. On peut supposer que ces heures correspondent à une période où il dort, compte tenu de l'activité réduite pendant cette période.

Maintenant, intéressons-nous à la question de savoir si **les tweets supprimés par Donald Trump suivent une tendance similaire à celle de ses tweets publiés**.

Tendance des tweets supprimés par heure



L'analyse de l'histogramme des tweets supprimés révèle des tendances intéressantes. La majorité écrasante des tweets supprimés se concentre autour de midi, avec **un pic de plus de 150 tweets supprimés**. Une autre concentration notable se situe pendant les heures nocturnes, ce qui pourrait suggérer une activité réduite de l'équipe de communication. En contraste, les heures suivant **l'après-midi ont très peu d'activité de suppression**, ne dépassant que **rarement un total de 20 tweets**.

Sur la base de ces observations, on peut se demander si l'équipe de communication est principalement active après midi (de 13h à minuit) afin de minimiser la suppression des tweets. De même, on peut déduire que Donald Trump lui-même est principalement actif à midi et pendant les heures nocturnes.

Warning : Cependant, il est important de noter que ces déductions sont basées sur des observations générales et qu'il est difficile de tirer des conclusions définitives sans une analyse plus approfondie et d'informations supplémentaires.

Les tendances des mots utilisés par Donald Trump

Dans le cadre de cette étude statistique, j'ai choisi de me concentrer sur certains mots clés spécifiques dans les tweets de Donald Trump, à savoir : **“COVID”, “CNN”, “War”, “Hillary Clinton”, “Fake News”, “Putin” et “Joe Biden”**. J'ai fait ce choix en considérant que ces mots font partie intégrante de sa vie, de sa communication générale et des événements marquants liés à sa carrière politique. Mon objectif principal est d'analyser si certains de ces mots ont été censurés plus fréquemment que d'autres et, éventuellement, de comprendre les raisons sous-jacentes à cette censure. En étudiant les tendances de censure, nous pourrions obtenir des informations précieuses sur les éventuelles restrictions imposées aux discours de Donald Trump et les motifs qui pourraient expliquer cette censure sélective.

Nombre total de tweets par mot clé

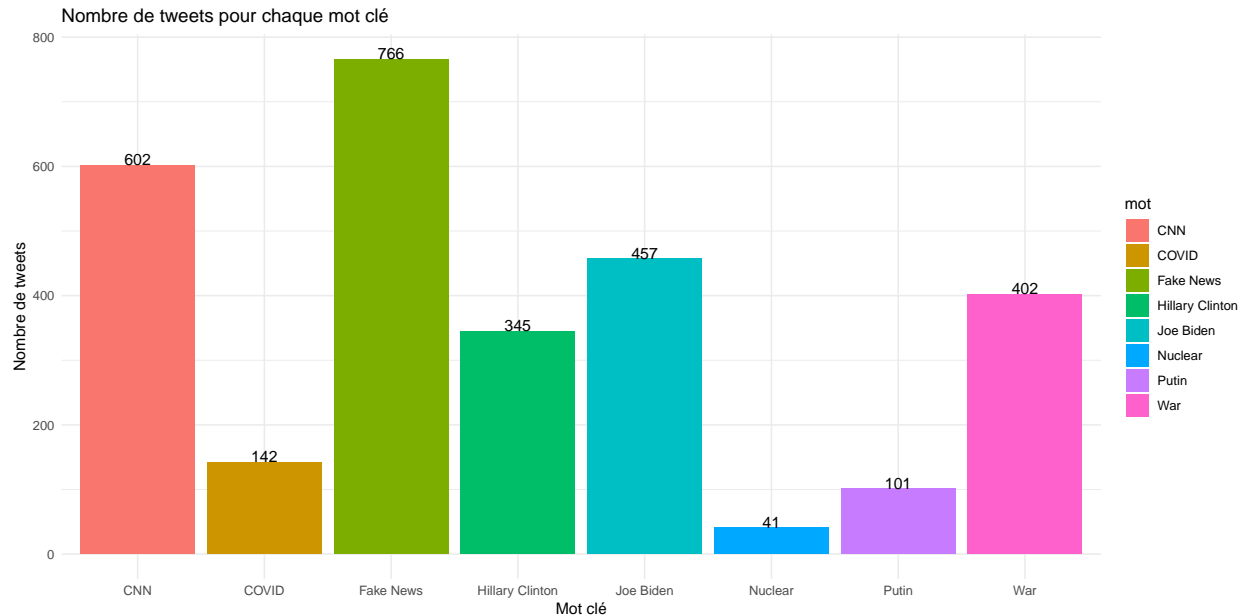
Questions: Y a-t-il une tendance observée dans l'utilisation des mots par Donald Trump ? Certains de ces mots ont-ils fait l'objet d'une censure particulière ?

Individu: un mot clé dans l'ensemble des tweets de Donald Trump

Population: L'ensemble des tweets de Donald Trump de 2009-2017.

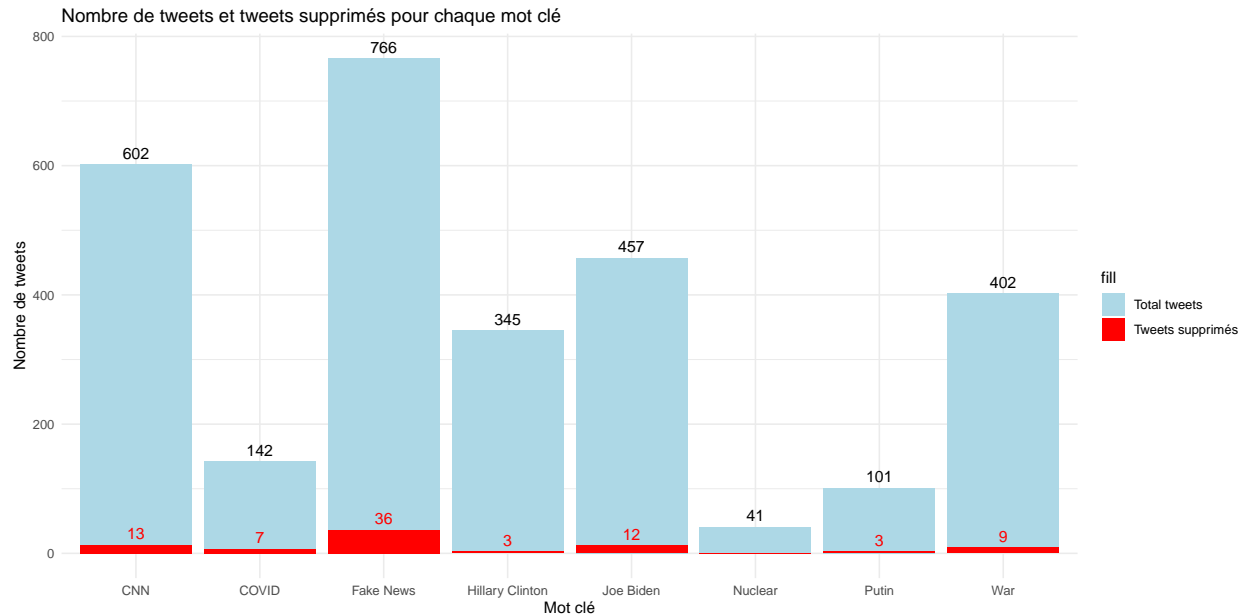
Échantillon: Kaggle Dataset Donald Trump tweets de 2009-2021

Variable && modalité: Nombre de tweets par mot clé, Nombre tweets supprimé par mot clé, quantitative discrète



Fidèle à ce dont on aurait pu penser, le terme **“Fake News”** se positionne en tête avec **766 tweets**, révélant ainsi l'utilisation de cette expression comme stratégie de communication de Donald Trump. En accord avec ses stratégies provocatrices envers certains adversaires politiques, les termes **“Hillary Clinton”**, **“Joe Biden”** et **“CNN”** - considéré comme un média hostile à Donald Trump - sont également largement représentés, avec respectivement **345 (HC)**, **457 (JB)** et **602 tweets (CNN)**. Cette observation confirme la tendance de Donald Trump à attaquer ces opposants politique. Cependant, il reste à déterminer si certains de ces mots clés, qui se démarquent par leur fréquence, ont été davantage sujets à la suppression de tweets que d'autres.

Nombre total de tweets supprimés et postés par mot clé



Ce diagramme présente plusieurs observations pertinentes, mais l'une d'entre elles retient particulièrement l'attention. Malgré le fait que le terme “fake news” ait été le mot clé le plus représenté dans notre liste de mots, il est également celui qui a été proportionnellement le plus supprimé, avec plus de **36 tweets** contenant ce terme ayant été retirés.

Cette constatation soulève plusieurs questions cruciales quant à l'interprétation objective de ces données et au rôle de Twitter en tant que garant de la vérité absolue. Peut-on considérer que Twitter, en supprimant ces tweets contenant le terme “fake news”, détient une autorité supérieure dans la détermination de la véracité des informations ?

De plus, cela soulève également des interrogations sur la place de la démocratie dans la gestion de l'information et le rôle des réseaux sociaux dans la prise de décision concernant la véracité des informations. En fin de compte, la question est de savoir **si Twitter est réellement neutre dans la gestion de l'information qu'il relaye**, et si un réseau social a plus d'autorité qu'un président pour décider si une information est juste ou non.

Utilisation de la liste de mots-clés par année

Questions: Y a-t-il une tendance observée dans l'utilisation des mots-clés de la liste pour chaque année ?

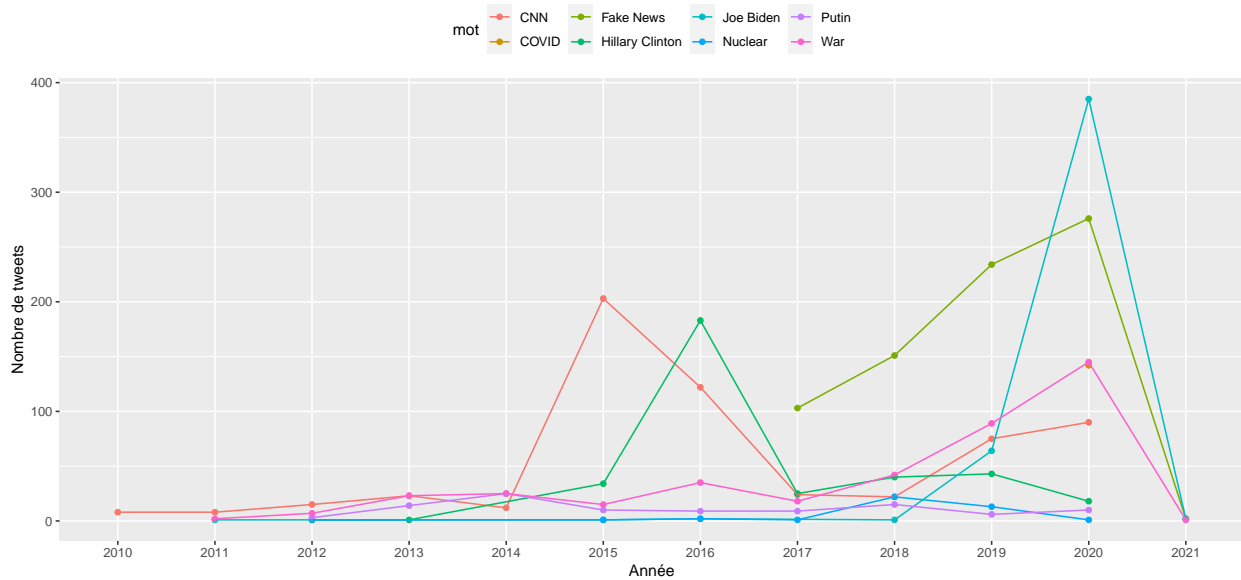
Individu: un mot clé de la liste l'ensemble des tweets de Donald Trump

Population: L'ensemble des tweets de Donald Trump de 2009-2017.

Échantillon: Kaggle Dataset Donald Trump tweets de 2009-2021

Variable && modalité: Nombre de tweets par mot-clé par année, quantitative discrète

Fréquence d'utilisation des mots clés par année



L'analyse du diagramme en nuage de points représentant les mots clés utilisés par année dans les tweets de Donald Trump permet de confirmer plusieurs observations significatives. Pendant les périodes électorales, Donald Trump adopte une approche agressive en **mentionnant spécifiquement ses opposants**, tels que Joe Biden en 2020, ainsi que CNN et Hillary Clinton en 2015 et 2016. Chaque année semble correspondre à une cible spécifique : **CNN en 2015**, **Hillary Clinton en 2016**, et **Joe Biden en 2020**.

Par ailleurs, une autre observation intéressante concerne l'émergence du terme "**fake news**", qui n'était **jamais utilisé avant 2017**. Cette évolution peut être attribuée à un événement particulier dans la vie de Donald Trump, et ce terme est devenu un élément clé de sa communication politique. Il convient de se demander ce qui aurait pu expliquer l'apparition de ce terme en 2017, tel que le départ, le changement, de sa boîte de communication. Notons également que c'est à partir de cette année que le nombre de tweets de Donald Trump a commencé à augmenter, en parallèle avec l'utilisation croissante du terme "fake news" (voir le diagramme du nombre de tweets par année).