

Comprendre l'influence de Donald Trump sur Twitter : Quels étaient les secrets de sa stratégie sur les réseaux sociaux ?

Objectif

Qu'est-ce qui a rendu la stratégie Twitter de Donald Trump si puissante ?

Ce projet explore et analyse les facteurs derrière le succès et l'influence de Donald Trump sur Twitter. Vous devrez investiguer :

1. Les stratégies qui ont généré un fort engagement du public (favoris et retweets).
2. Les thèmes et sujets qui ont le plus résonné auprès de son audience.
3. L'évolution de son style de communication dans le temps et le lien entre ce style de communication et des événements politiques ou sociaux majeurs.

La Question Centrale

Comment Donald Trump a-t-il utilisé Twitter pour influencer l'opinion publique, engager son audience et passer d'une personne que peu connaissent à Président des USA ?

Description du Projet

Dans ce projet, vous allez analyser un dataset contenant les tweets de Donald Trump et les métadonnées associées. Vous devrez :

- explorer les thèmes et sentiments clés dans ses Tweets afin de comprendre sa stratégie de communication.
- Analyser les tendances d'engagement pour identifier les Tweets ayant suscité le plus d'attention.
- Corréler le contenu des tweets avec les grands événements politiques ou sociaux pour examiner ce qui a fait leur impact.
- Analyser le rôle des habitudes de publication (temps de la journée, appareils utilisées, etc.) dans l'amplification de la portée des tweets.

Blocs Attendus :

1. Analyse de l'Engagement :

- Identifier et visualiser quels types de tweets (par exemple : tweets originaux, retweets ou sujets spécifiques) ont généré le plus d'engagement.

2. Analyse du Contenu :

- Identifier les thèmes récurrents, les schémas linguistiques et les phrases clés qui définissent son style de communication.

3. Analyse Temporelle :

- Explorer comment le comportement de tweet de Trump et la réponse du public ont évolué dans le temps, en particulier lors des événements majeurs.

4. Insights Stratégiques :

- Fournir un résumé des éléments qui ont rendu la stratégie Twitter de Trump efficace et des leçons applicables à d'autres campagnes sur les réseaux sociaux.

Tâches :

1. Prétraitement

- **Objectif** : Nettoyer et préparer le dataset pour l'analyse.
 - **Tâches** :
 - Gérer les données manquantes ou incohérentes.
 - Comprendre les spécificités du dataset :
 - Longueur des tweets,
 - Mots clés les plus utilisés à travers le temps,
 - Utilisation de liens/mentions/hashtags,
 - Le ratio tweet/retweet,
 - La densité des tweets/retweets à travers le temps.
 - Choisir une approche de nettoyage des données :
 - Comment traiter les liens/mentions/hashtags ?
 - Comment nettoyer les tweets (textes courts) ?
 - Comment traiter les NER / POS ?
 - Devriez-vous diviser le dataset selon les périodes, le nombre de favoris, le nombre de retweets, etc. ?
 - Choisir une approche de vectorisation appropriée aux analyses qui vont suivre :
 - N-gram, tfidf, word2vec ?
-

2. Modèles d'Engagement

- **Objectif** : Découvrir ce qui génère de l'engagement (favoris et retweets).
- **Tâches** :
 - Analyser les différences d'engagement entre :

- Tweets et retweets,
 - Appareils (par exemple : Twitter Media Studio, iPhone, Android, etc.),
 - Heure de la journée,
 - Sentiments générés,
 - Etc.
- Identifier les tweets/retweets les plus engageants et expliquer leur succès selon les variables citées ci-dessus ainsi que les événements politiques/sociaux qui peuvent contextualiser ces tweets/retweets.
-

3. Analyse Thématique

- **Objectif** : Explorer les thèmes récurrents et les sujets clés dans les tweets de Trump.
 - **Tâches** :
 - Effectuer une **analyse de fréquence des mots** pour trouver les mots couramment utilisés.
 - Utiliser une **analyse des n-grams** pour identifier les phrases populaires.
 - Appliquer la **modélisation de thèmes (LDA)** pour découvrir des sujets sous-jacents.
-

4. Sentiments et Ton Émotionnel

- **Objectif** : Comprendre le ton émotionnel des tweets de Trump et son impact.
 - **Tâches** :
 - Appliquer une analyse de sentiments pour classifier les tweets en positifs, négatifs ou neutres.
 - Analyser comment le ton émotionnel a varié dans le temps ou à travers les thèmes.
-

5. Tendances Temporelles

- **Objectif** : Examiner comment la stratégie Twitter de Trump a évolué dans le temps.
- **Tâches** :
 - Grouper les tweets par mois/année et analyser les changements dans :
 - La fréquence des tweets.
 - Les métriques d'engagement.
 - Les thèmes de contenu.

- Corréler les pics d'activité avec des événements clés (par exemple : élections, grandes annonces).

Livrables attendus

1. Rapports :

- Résumez toutes les analyses dans un notebook Jupyter organisé.
- Incluez des visualisations claires (graphiques temporels, nuages de mots, etc.).

2. Recommandations stratégiques:

- Proposer des insights sur des stratégies Twitter efficaces basées sur l'analyse.

3. Une Présentation finale :

- Destinée à un public non-technique, la présentation doit se concentrer sur :
 - L'analyse de l'engagement,
 - L'analyse du contenu,
 - L'analyse temporelle,
 - Les insights stratégiques.
- Cette présentation est destinée à des profanes en programmation, en analyse des données et en Text Mining. L'objectif est de partager les résultats de l'analyse et non le processus de l'analyse en soit.

4. Pour chaque personne :

- Votre conversation avec votre IA générative
- Une réflexion personnelle honnête sur ce projet de groupe :
 - 1-2 paragraphes maximum.
 - Comment avez-vous vécu ce travail de groupe ?
 - Comment GenAI s'est elle intégré dans votre travail de groupe ?

Le dataset :

- Les tweets de trump du 05/04/2009 au 01/08/2021 :
<https://huggingface.co/datasets/fschlatt/trump-tweets>
- Le dataset est téléchargeable sous format .parquet.

La description du dataset :

- <https://www.kaggle.com/datasets/headsortails/trump-twitter-archive>
- <https://www.thetrumparchive.com/>