**DR False**

**Produit False**

**\*S-produit True\***

**Problème False**

**\*S-problème True\***

**Desc-plainte-C True**

**\*RéP-l'entreprise True\***

**Société False**

**\*État True\***

**\*Code postal True\***

**Mots-clés True\***

**\*Consentement du consommateur True\***

**Soumis via False**

**Date d'envoi à l'entreprise False**

**Rép-consommateur False**

**Rép-délais ? False**

**\*Contesté par le consommateur True\***

**ID-plainte False**

**Desc-plainte-C**

Dépend de la colonne produit

**RéP-l'entreprise**

Dépend de la colonne société et problème , produit , soumis via

**Consentement du consommateur**

Dépend de la colonne rep entreprise et problème

|  |  |  |  |
| --- | --- | --- | --- |
| Consent not provided | Consent with drawn | Other | Consent provided |

**Contesté par le consommateur**

|  |  |
| --- | --- |
| Yes | No |

Dépend de la colonne societe et rep entreprise

**Contesté par le consommateur**

1. **S-produit et S-problème :**
   * Ces deux colonnes peuvent contenir des informations sur la catégorie spécifique de produit et le problème associé signalé par le consommateur. Ils pourraient être des prédicteurs importants pour déterminer si un consommateur conteste ou non.
2. **Desc-plainte-C** :
   * Cette colonne pourrait contenir des descriptions plus détaillées des plaintes des consommateurs, ce qui pourrait également être utile pour prédire si elles sont contestées ou non.
3. **RéP-l'entreprise**, **État**, **Code postal**, **Mots-clés**, **Consentement du consommateur** :
   * Ces colonnes peuvent également contenir des informations pertinentes pour la prédiction. Par exemple, l'état et le code postal pourraient indiquer des variations géographiques dans les comportements de contestation, tandis que les mots-clés et le consentement du consommateur pourraient donner des indications sur les facteurs qui pourraient influencer la décision de contester.
4. **Soumis via :**
   * Bien que cette colonne soit marquée comme "False" dans votre liste, elle pourrait également être informative si elle était disponible. Elle pourrait indiquer la source à travers laquelle la plainte a été soumise, ce qui pourrait avoir un impact sur la probabilité de contestation.

l'utilisation simultanée de "Soumis via", "S-produit" et "S-problème" dans votre analyse peut constituer une approche robuste et complète pour prédire les valeurs manquantes dans la colonne "Contesté par le consommateur" de votre ensemble de données

**Consentement du consommateur**

Pour prédire les valeurs manquantes dans la colonne "Consentement du consommateur" en utilisant d'autres variables de votre ensemble de données, vous pouvez envisager plusieurs approches. Voici quelques suggestions sur les colonnes avec lesquelles vous pourriez travailler :

1. **Problème et S-problème** :
   * Il est possible que le type de problème signalé par le consommateur puisse influencer leur décision de fournir ou de ne pas fournir leur consentement. En examinant la relation entre le problème signalé et le consentement du consommateur, vous pourriez identifier des tendances ou des corrélations significatives.
2. **Produit et S-produit** :
   * De manière similaire, le type de produit concerné par la plainte pourrait également jouer un rôle dans la décision de consentement du consommateur. En explorant cette relation, vous pourriez découvrir des schémas intéressants qui pourraient être utiles pour prédire les valeurs manquantes.
3. **Soumis via** :
   * Le canal par lequel la plainte a été soumise pourrait également être pertinent. Par exemple, les consommateurs pourraient être plus enclins à donner leur consentement lorsqu'ils utilisent certains canaux de communication spécifiques.
4. **RéP-l'entreprise** :
   * La manière dont l'entreprise a répondu à la plainte pourrait également influencer la décision de consentement du consommateur. Par exemple, un service client efficace et réactif pourrait augmenter la probabilité de consentement.
5. **Code postal ou État** :
   * Les caractéristiques géographiques telles que le code postal ou l'état pourraient également jouer un rôle dans la décision de consentement du consommateur, en fonction de divers facteurs contextuels.
6. **Mots-clés** :
   * Les mots-clés utilisés dans la description de la plainte pourraient également être informatifs. Certains termes pourraient indiquer un niveau de satisfaction plus élevé ou des préoccupations particulières qui pourraient influencer la décision de consentement.

Haut du formulaire

Nous avons choisi d'utiliser les variables "Soumis via" et "RéP-l'entreprise" pour prédire les valeurs manquantes dans la colonne "Consentement du consommateur". Ces variables sont cruciales car elles captent respectivement le canal de soumission des plaintes par les consommateurs et la réponse de l'entreprise à ces plaintes. En explorant leur relation avec la variable "Consentement du consommateur", nous cherchons à identifier des tendances significatives qui pourraient influencer la propension des consommateurs à donner leur consentement.

**Desc-plainte-C**

1. **Problème et S-problème** :
   * Les problèmes signalés par les consommateurs pourraient être étroitement liés à la description de la plainte. En examinant la relation entre les problèmes signalés et la description de la plainte, vous pourriez identifier des tendances ou des corrélations significatives qui pourraient être utilisées pour prédire la description de la plainte.
2. **Produit et S-produit** :
   * De manière similaire, le type de produit concerné par la plainte pourrait également être lié à la description de la plainte. En explorant cette relation, vous pourriez découvrir des schémas ou des motifs récurrents dans les descriptions de plainte pour des produits particuliers.
3. **Mots-clés** :
   * Les mots-clés utilisés dans la description de la plainte pourraient également être informatifs pour prédire la description de la plainte. Certains mots ou phrases pourraient être associés à des types spécifiques de plaintes ou à des préoccupations particulières des consommateurs.
4. **RéP-l'entreprise** :
   * La manière dont l'entreprise a répondu à la plainte pourrait également influencer la description de la plainte. Par exemple, une réponse insatisfaisante de l'entreprise pourrait conduire à des descriptions de plainte plus détaillées ou plus négatives.

**RéP-l'entreprise**

1. **Problème et Produit** :
   * Les caractéristiques du problème signalé par le consommateur et du produit concerné pourraient être liées à la réponse de l'entreprise. Par exemple, certaines réponses de l'entreprise pourraient être plus fréquentes pour des types spécifiques de problèmes ou de produits.
2. **Soumis via** :
   * Le canal par lequel la plainte a été soumise pourrait également influencer la réponse de l'entreprise. Par exemple, l'entreprise pourrait avoir des protocoles ou des politiques différentes pour répondre aux plaintes soumises par téléphone par rapport à celles soumises par e-mail.
3. **Consentement du consommateur** :
   * La décision du consommateur de consentir ou non à une réponse de l'entreprise pourrait également être liée à la réponse de l'entreprise elle-même. Par exemple, une réponse satisfaisante de l'entreprise pourrait augmenter la probabilité que le consommateur donne son consentement à la réponse.
4. **Desc-plainte-C** :
   * La description de la plainte des consommateurs pourrait également fournir des indications sur la réponse de l'entreprise. Par exemple, certaines descriptions de plainte pourraient nécessiter des réponses spécifiques de la part de l'entreprise.
5. **Société** :
   * La variable "Société" représente l'entité ou la division spécifique de l'entreprise chargée de traiter les plaintes des consommateurs. En examinant les différentes sociétés impliquées dans ce processus, vous pourriez identifier des schémas ou des tendances qui pourraient influencer les réponses fournies aux consommateurs.

Dans notre analyse pour prédire les réponses de l'entreprise dans la colonne "RéP-l'entreprise", nous avons pris en compte plusieurs variables, notamment "Société", "Problème", "Produit" et "Soumis via". Nous avons choisi de nous concentrer sur la colonne "Société" en raison de son importance potentielle dans le processus de traitement des plaintes des consommateurs. En examinant les différentes entités ou divisions de l'entreprise impliquées dans ce processus, nous avons cherché à identifier des schémas ou des tendances qui pourraient influencer les réponses fournies aux consommateurs.

Nous avons également inclus les variables "Problème", "Produit" et "Soumis via" dans notre analyse, car elles offrent des informations supplémentaires sur le contexte spécifique de chaque plainte. Nous pensons que ces variables pourraient être étroitement liées aux réponses de l'entreprise et pourraient donc être utiles pour prédire ces réponses. Par exemple, certaines catégories de produits ou de problèmes pourraient nécessiter des réponses spécifiques de la part de l'entreprise, tandis que le canal par lequel la plainte a été soumise pourrait également influencer la manière dont elle est traitée.

**Mot clé**

Pour remplir la colonne "Mots-clés" de votre dataset, vous pouvez utiliser différentes colonnes comme source d'information pour extraire des mots-clés pertinents. Voici quelques suggestions sur les colonnes que vous pourriez utiliser comme partie de votre xtrain (ensemble d'entraînement) pour cette tâche :

* Produit, S-produit et S-problème : Ces colonnes peuvent contenir des informations sur le type de produit et le type de problème rencontré par le consommateur. Vous pouvez extraire des mots-clés à partir de ces colonnes pour décrire le produit et le problème de manière succincte.
* Desc-plainte-C : Cette colonne contient probablement une description détaillée de la plainte du consommateur. Vous pouvez utiliser des techniques de traitement du langage naturel (NLP) pour extraire des mots-clés significatifs à partir de ces descriptions.
* RéP-l'entreprise : Cette colonne pourrait contenir des réponses de l'entreprise aux plaintes des consommateurs. Les mots-clés dans ces réponses pourraient également être utiles pour comprendre le contexte de la plainte.
* État et Code postal : Si ces colonnes contiennent des informations sur la localisation géographique des plaintes, vous pourriez utiliser des mots-clés liés à la région ou à la localisation pour enrichir votre ensemble de mots-clés.
* Contesté par le consommateur : Cette colonne pourrait indiquer si la plainte a été contestée par le consommateur. Les mots-clés associés à cette information pourraient également être pertinents pour comprendre le contexte de la plainte.

En utilisant une combinaison de ces colonnes et en appliquant des techniques de traitement du langage naturel appropriées, vous pouvez créer un ensemble de mots-clés qui capture efficacement les aspects pertinents des plaintes des consommateurs.