

# Projet

## Contexte

Dans le domaine de l'aviation, la satisfaction client est un indicateur crucial qui influence directement la fidélité des clients et la rentabilité. Afin de mieux comprendre les facteurs qui contribuent à la satisfaction des voyageurs, vous travaillerez sur un ensemble de données qui recense les réponses des clients suite à leur expérience de vol. L'objectif est d'identifier des segments de clients homogènes en fonction de leurs réponses pour améliorer les services proposés.

## Données

Le dataset fourni est un fichier CSV contenant plusieurs variables relatives à l'expérience de voyage des clients. Les variables incluent, entre autres, des informations démographiques, le type de voyage, la classe de voyage, la distance de vol, et des évaluations de différents aspects du service (confort des sièges, divertissement en vol, service client, etc.).

## Objectifs

**Prétraitement des données :** Commencez par nettoyer et préparer les données pour l'analyse. Cela peut inclure la gestion des valeurs manquantes, la normalisation des variables numériques, et le codage des variables catégorielles.

**Sélection des variables :** Sélectionnez les variables pertinentes pour l'analyse de clustering. Vous devrez justifier le choix de ces variables dans votre rapport. Considérez l'impact potentiel de chaque variable sur la satisfaction client et la faisabilité technique de leur utilisation dans un modèle SOM.

**Clustering avec SOM (Self-Organizing Maps) :** Utilisez l'algorithme SOM pour effectuer un clustering sur l'ensemble de données prétraité. SOM vous aidera à identifier des groupes de clients ayant des profils de satisfaction similaires. Assurez-vous de choisir des paramètres appropriés pour le SOM en fonction de la taille de votre dataset et de la complexité des données.

**Analyse des clusters :** Une fois les clusters formés, analysez-les pour comprendre les caractéristiques communes des clients au sein de chaque groupe. Recherchez des tendances, des préférences, et des domaines d'amélioration potentielle en matière de services offerts.

**Rapport :** Rédigez un rapport détaillé de votre analyse. Le rapport doit inclure une introduction au projet, une description des données, la méthodologie employée pour le prétraitement des données et le clustering, une discussion sur les résultats obtenus, ainsi que des recommandations basées sur l'analyse des clusters.

## Livrables

Code source complet de votre projet (incluant le nettoyage des données, la sélection des variables, l'implémentation du SOM, et l'analyse des clusters).

Un rapport d'analyse en PDF détaillant les étapes réalisées, les choix méthodologiques, les résultats obtenus, et les recommandations pour améliorer la satisfaction client basées sur l'analyse des données.

## **Variables**

**id** : Un identifiant unique pour chaque réponse ou participant à l'enquête.

**Gender** : Le genre du passager.

**Customer Type** : Le type de client, pouvant indiquer un voyageur fréquent (disloyal customer) ou un voyageur occasionnel (casual traveler).

**Age** : L'âge du passager.

**Type of Travel** : Le type de voyage, indiquant s'il s'agit d'un voyage d'affaires ou personnel.

**Class** : La classe de cabine dans laquelle le passager a voyagé.

**Flight Distance** : La distance du vol en kilomètres.

**Inflight wifi service** : L'évaluation du service wifi à bord (échelle de notation).

**Departure/Arrival time convenient** : L'évaluation de la commodité des horaires de départ et d'arrivée (échelle de notation).

**Ease of Online booking** : La facilité de réservation en ligne (échelle de notation).

**Gate location** : La satisfaction concernant l'emplacement de la porte d'embarquement (échelle de notation).

**Food and drink** : L'évaluation de la qualité de la nourriture et des boissons à bord (échelle de notation).

**Online boarding** : La satisfaction avec le processus d'embarquement en ligne (échelle de notation).

**Seat comfort** : Le confort du siège (échelle de notation).

**Inflight entertainment** : La qualité des divertissements en vol (films, TV, musique, jeux) (échelle de notation).

**On-board service** : L'évaluation du service à bord par le personnel de cabine (échelle de notation).

**Leg room service** : La satisfaction concernant l'espace disponible pour les jambes (échelle de notation).

**Baggage handling** : L'efficacité du traitement des bagages (échelle de notation).

**Checkin service** : La satisfaction avec le service d'enregistrement (échelle de notation).

**Inflight service** : L'évaluation du service en vol, à l'exception de celui fourni par le personnel de cabine (échelle de notation).

**Cleanliness** : La propreté de l'avion (échelle de notation).

**Departure Delay in Minutes** : Le retard de départ en minutes.

**Arrival Delay in Minutes** : Le retard à l'arrivée en minutes.

**satisfaction** : L'indicateur de satisfaction globale du passager, pouvant être positif (satisfait) ou négatif (insatisfait).