

CALCULER UN SCORE ROBUSTE ET ÉQUITABLE

Partie Client :

- Numéro d'identifiant qui permettra de reconnaître chaque fois que le client à poster une note

Partie Chauffeur :

- Numéro d'identifiant permettant de reconnaître le chauffeur
- Un score donné par le client
- Un numéro de rang calculé grâce à toutes les notes qu'il a obtenu
- Le nombre de note qui lui sont attribué

Objectif :

- Calculer une moyenne de score donnée au chauffeur par les clients de manière équivalente
- Ce calcul permettra de faire un classement selon différents critères que l'on jugera nécessaire

Cas particulier :

Si l'identifiant d'un client apparaît plusieurs fois, on calcul d'abord la moyenne de toutes les notes qu'il a mis au chauffeur. Puis cette valeur aura le même poids que pour un client aurait noté qu'une seule fois (chaque client à le poids de 1 dans la moyenne d'un chauffeur).

Ce système permet d'avoir une moyenne équitable. Dans le cas où un client met que des bonnes notes et on s'aperçoit que les autres clients déposent des notes assez faibles, ce moyen permet de se poser des questions sur le chauffeur.

La variance est aussi calculée ce qui permet de voir la dispersion des notes obtenues par le chauffeur et donc de le classer si les moyennes sont semblables.

La note est calculée en fonction :

- Du **volume d'avis reçus** (Nombre de notes)
- De la **fraîcheur** (s'ils sont récents ou anciens)
- De la **qualité des avis** : la note globale attribuée par les voyageurs

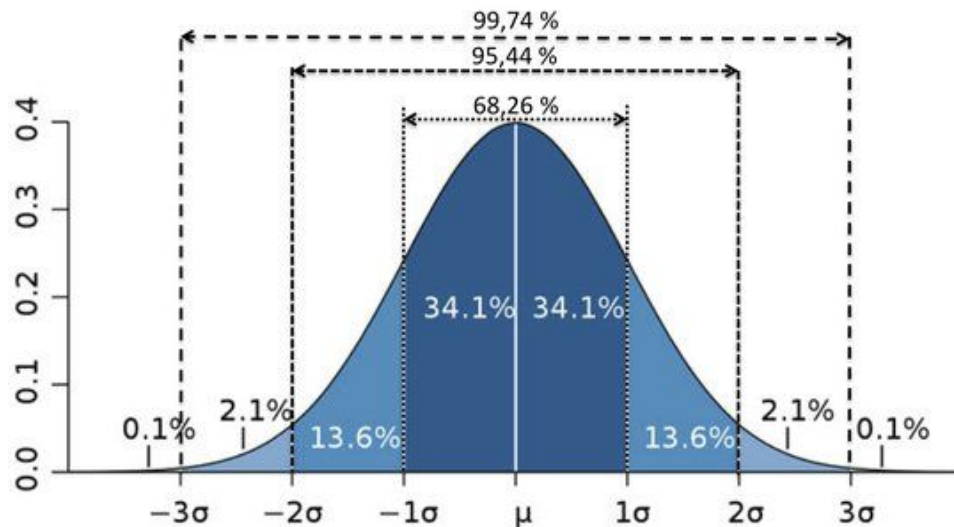
La note du chauffeur est donc calculé en fonction de la date des notes déposées, du nombre de notes que le chauffeur détient, ainsi que la variance des notes de chaque utilisateurs.

Idées :

- On peut aussi noter la moyenne en fonction du temps de la note déposée : une faible note récemment déposée sera plus importante qu'une forte note déposée il y a plus longtemps (temps à définir)

Pour cela utiliser la formule suivante : $A_i = \alpha * Valeur + (1 - \alpha)A(i - 1)$

- On va s'appuyer sur la courbe de Gauss, avec la loi normale pour vérifier la distribution. C'est à dire que nous centralisons les données, quitte à perdre certaines données éloignées de la moyenne. On retient toutes les données comprises dans l'intervalle de 95 % ($[m - 1,96 \sigma ; m + 1,96 \sigma]$). Soit on effectue les calculs avec une marge d'erreur de 10%.



- On ne peut pas utiliser les mêmes algorithmes si le chauffeur a une seule note pour le classer car certains algorithmes comme la loi Normale qui nécessitent plusieurs valeurs pour être définies.
- Une autre possibilité est d'utiliser l'algorithme des moindres carrés. Pour rentrer dans les détails :
 - Cette méthode, une des plus précises, consiste à déterminer la droite théorique dont les coordonnées sont la moyenne arithmétique de toutes les données.
 - Pour faire plus simple, c'est la droite qui passe au plus près de tous les points de données.
 - Cette droite se nomme la droite de régression.

Elle permet de faire du prévisionnel à partir de données existantes.