**Rotation du personnel ( donnée turnover.csv)**

Ensemble de données sur la rotation du personnel utilisé à l'origine pour un modèle d'analyse de survie

**À propos de l'ensemble de données :**

Non, il ne s’agit pas de survivre à la noyade ou quelque chose comme ça (juste pour illustrer).

Cet ensemble de données sur la rotation du personnel est un ensemble de données réel partagé à partir du [blog d'Edward Babushkin](https://edwvb.blogspot.com/2017/10/employee-turnover-how-to-predict-individual-risks-of-quitting.html) utilisé pour prédire le risque de démission d'un employé (avec un modèle d'analyse de survie). Edward Babushkin a expliqué que « l'analyse de survie est l'un des algorithmes les plus importants, mais ce n'est pas l'algorithme le plus populaire pour prédire la rotation du personnel. Les analystes utilisent des algorithmes plus familiers comme la régression logistique mais, par exemple, Pasha Roberts écrit : « N'utilisez pas de méthodes logistiques pour prédire l'attrition ! ». Je pense que nous ne pouvons appliquer cette méthode que pour une situation à court terme, par exemple si l'employé a travaillé plus ou moins de trois mois. Si notre objectif est de prédire les risques de démission individuels, alors la meilleure méthode est l'analyse de survie. »

stag : Le temps jusqu’à la démission (en moi)

event : vaut 1 si l’individu est censuré, 0 sinon

gender : sexe (f = femme, m = homme)

age : l’âge de la personne

industry : type d’entreprise

profession :

coach : Présence d'un coach (formation) en probation

head\_gender : sexe du chef de l’employé

greywave : white si le travail est déclaré, grey sinon

way : moyen de transport

extraversion : score d’extraversion

independ : score d’indépendance

selfcontrol : score de self contrôle

anxiety : score d’anxiété

novator : score de novateur/novatrice

**Inspirations:**

Cet ensemble de données peut être utilisé pour prédire le taux de rotation du personnel, l'analyse de survie des employés, la modélisation de l'augmentation ou même l'analyse de survie de l'augmentation.

<https://www.kaggle.com/datasets/davinwijaya/employee-turnover?resource=download>

**À propos de Dataset (HR-Employee-Attrition.csv)**

L'ensemble de données contient les mêmes fonctionnalités que l'ensemble de données IBM, avec quelques fonctionnalités supplémentaires pour nous aider dans le projet, qui comprend l'analyse de la survie et la prédiction d'un employé, qu'il soit attrite ou non.

<https://www.kaggle.com/datasets/rv1922/employee-data-attrition>

**About Dataset (Hunger Games survival analysis data set.csv)**

1. Pronostiquez quels tributs auront le temps de survie le plus long dans les futurs jeux de la faim.
2. Construisez un modèle pour prédire comment différentes covariables affectent les chances de gagner ou d'être dans une alliance.
3. Classifiez avec un modèle qui peut prédire si un hommage sera un hommage à la carrière ou non

Cet ensemble de données contient des informations sur tous les hommages qui ont participé aux 74e Hunger Games. Serez-vous en mesure d'utiliser ces données pour prédire qui survivra ?

1. name : Le nom de l'hommage. (Chaîne)
2. district : Le quartier de l'hommage. (Chaîne)
3. sexe : Le sexe de l'hommage. (Chaîne)
4. age : L'âge de l'hommage. (Entier)
5. bénévole : si l'hommage a été bénévole pour les Hunger Games. (Booléen)
6. carrière : si l'hommage était un hommage à la carrière. (Booléen)
7. rating : La note de l'hommage. (Flotteur)
8. rating\_ave : La note moyenne de l'hommage. (Flotteur)
9. rating\_rand : La note aléatoire de l'hommage. (Flotteur)
10. survival\_days : Le nombre de jours pendant lesquels l'hommage a survécu. (Entier)
11. rang : le dernier rang du tribut. (Entier)
12. gagnant : Si l'hommage a remporté les Hunger Games. (Booléen)
13. alliance : l'alliance finale du tribut. (Chaîne)

<https://www.kaggle.com/datasets/thedevastator/the-hunger-games-dataset-a-survival-analysis>

<https://lbelzile.github.io/math60602/survie.html>

Cox Boosting

<https://link.springer.com/article/10.1007/s00180-015-0642-2>

<https://link.springer.com/article/10.1007/s00180-017-0773-8>

<https://rdrr.io/cran/CoxBoost/man/CoxBoost.html>

Forêt aléatoire de survie

<https://depot-e.uqtr.ca/id/eprint/11471/1/eprint11471.pdf>

<https://cran.r-project.org/web/packages/ggRandomForests/vignettes/ggrfRegression.html>

<https://www.randomforestsrc.org/articles/survival.html>

<https://www.randomforestsrc.org/articles/getstarted.html>

Pour la survie,la statistique des nœuds terminaux le TNS est l'estimateur de Kaplan-Meier et la fonction de risque cumulé de Nelson-Aalen

Erreur de prédiction de survie est basé sur l'indice C de Harrell (1 moins la concordance) en utilisant la mortalité à des fins de comparaison