

# Projet d'analyse de survie

2017

L'objectif du projet est de vous faire appliquer les méthodes d'analyse de survie sur des données réelles. La première partie du projet consiste à trouver un jeu de données adapté à cette étude. Il peut s'agir de données publiques (Kaggle, OpenData, data.gouv.fr, ...) qui n'ont pas déjà été analysées de la sorte, ou de données issues de votre entreprise. Les projets sont à réaliser par groupe de 2.

Avant le **16 avril 2017** à minuit, vous devez m'envoyer par mail (thibault.allart@gmail.com) :

- un lien vers les données
- un descriptif sommaire : variable d'intérêt, variables explicatives, censure

L'évaluation se fera sur la qualité de l'analyse et l'interprétation des résultats obtenus, mais aussi sur l'originalité du jeu de données et la qualité du code.

A rendre avant le **14 mai 2017** à minuit:

- Le rapport de votre analyse en maximum 10 pages au format pdf
- Le code R (éventuellement au format notebook) commenté et exécutable.

Exemple de contenu :

- Description et visualisation de données
- Choix du modèle et sélection des variables explicatives
- Interprétation : Quelle est l'influence des covariables sur le risque d'évènement ?
- Prédiction : Taux de risque instantané et courbe de survie prédits par le modèle pour un individu dont les caractéristiques sont données.

**15 mai 2017** : présentation des projets.

- Présentation à l'oral de 10 minutes + questions
- Slides au format pdf.

Remarques : Ce projet compte pour 50% de la note du module. Deux membres d'un binôme peuvent avoir des notes différentes selon leur implication dans le projet et la qualité de leur présentation.