

Exercices Pandas / Matplotlib

Question 1

Chargez les données du Titanic. Éliminez les colonnes avec des valeurs manquantes. Produisez la sous-matrice des personnes ayant payé leur billet une valeur inférieure à trois fois la valeur moyenne.

Dans toutes les questions suivantes on travaillera à partir de cette sous-matrice uniquement.

Question 2

Produisez un graphe des champs age et tarif en codant le genre avec une couleur. Que peut-on en déduire concernant le profil respectif des hommes et des femmes ?

Question 3

Produisez les 6 sous-matrices correspondant aux combinaisons des 3 classes et au 2 genres. Pour chacune de ces sous-matrices, effectuez un subplot affichant l'histogramme du taux de survie en fonction de l'âge. Commentez.

Question 4

Reprenez maintenant la matrice initiale et essayez de trouver d'autres variables impactant significativement le taux de survie, toutes choses égales par ailleurs.