

Master Pro M2, SITN, Université Claude Bernard, Lyon1

Analyse de données

année 2010-2011

PROJET(individuel)- ANALYSE DE DONNEES

à rendre au plus tard le 28 Février 2011

Pour chaque problème il faut d'abord faire une étude descriptive.

Problème 1. A l'adresse internet

<http://www.stats.ox.ac.uk/pub/PRNN/>

téléchargez le fichier "*fglass.dat*". La description des données vous la trouvez sur les pages suivantes. Il y a-t-il de types de glace qui se ressemblent? De quel point de vue? Il y a-t-il de types de glace atypiques? Pourquoi?

Problème 2. A la même adresse internet

<http://www.stats.ox.ac.uk/pub/PRNN/>

téléchargez le fichier "*Cushings.dat*". La description des données vous la trouvez sur les pages suivantes. Vous trouvez les mêmes données dans le package "MASS" du logiciel R., fichier "*Cushings*".

Peut-on prévoir les 3 types de syndrome de Cushing? Donnez pour les patients avec le type inconnu, les probabilités d'avoir chaque type de syndrome.

Problème 3. Dans le package "MASS" du logiciel R, téléchargez le fichier "*biopsy*". La description des données vous la trouvez sur les pages suivantes.

Peut-on prévoir le type de tumeur fonction des variables quantitatives mesurées?

Problème 4. Dans le package "MASS" du logiciel R, téléchargez le fichier "*farms*". La description des données vous la trouvez sur les pages suivantes.

Analyser ce tableau de contingence.

Problème 5. A venir